

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 08:06:58 ; Search time 67.41 Seconds
(without alignments)
1255.574 Million cell updates/sec

Title: us-09-673-302a-1

Perfect score: 4154
Sequence: 1 GPNICTRGVSSCQCCLAVS.....NNPLYKRNSTFNITRYRGT 762

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database: A_Geneseq_032802.*

1: /SID5/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID5/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID5/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID5/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID5/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID5/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID5/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID5/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID5/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID5/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID5/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID5/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID5/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID5/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID5/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID5/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID5/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID5/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID5/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	4119	99.2	788	20 AAY49553 Human endothelial
2	4119	99.2	788	20 AAY49567 Human antithrombin
3	4119	99.2	788	21 AAY92443 Wild type human GP
4	4119	99.2	788	21 AAY81461 Human integrin bet
5	4115	99.1	762	21 AAY92451 GPIIIa variant Arg
6	4115	99.1	788	15 AAR51500 Human platelet GPI
7	4114	99.0	762	21 AAY92456 GPIIIa variant Ser
8	4113	99.0	762	21 AAY92448 GPIIIa variant Leu
9	4113	99.0	762	21 AAY92450 GPIIIa variant Ser
10	4112	99.0	788	21 AAY92444 Variant human GPII
11	4111	99.0	762	21 AAY92452 GPIIIa variant Arg

12	4111	99.0	762	21 AAY92454
13	4111	99.0	762	21 AAY92455
14	4110	98.9	762	21 AAY92449
15	4108	98.9	762	21 AAY92453
16	3844	92.5	787	18 AAW13573
17	3768	90.7	718	11 AAR05936
18	3321	79.9	798	22 AAR07419
19	3282	79.0	720	18 AAW13574
20	2339	56.3	799	17 AAR02194
21	2156	51.9	788	13 AAR26332
22	2156	51.9	788	20 AAY43165
23	2150	51.8	788	22 AAB66747
24	1838	44.2	798	22 AAM93960
25	1831	44.1	798	22 AAB36937
26	1826	44.0	798	22 AAB84751
27	1776.5	42.8	805	22 AAM41146
28	1676	40.3	577	20 AAY43166
29	1636	39.4	557	13 AAR26333
30	1608	38.7	963	19 AAR70540
31	1537	37.0	845	22 AAB71129
32	1513.5	36.4	769	11 AAR07113
33	1513.5	36.4	769	16 AAR80108
34	1513.5	36.4	769	21 AAR803974
35	1513.5	36.4	793	22 AAB16399
36	1504.5	36.2	768	9 AAR80836
37	1494.5	36.0	769	13 AAR24256
38	1375	33.1	699	20 AAR81840
39	1373.5	33.1	565	21 AAB43718
40	1373.5	33.1	565	22 AAG3783
41	1343.5	32.3	676	18 AAR35853
42	1325	31.9	768	13 AAR27683
43	1304	31.4	768	13 AAR27684
44	1271	30.6	266	14 AAR39196
45	1265	30.5	266	14 AAR39195

ALIGNMENTS

RESULT 1
AAY49553
ID AAY49553 standard; Protein: 788 AA.
AC AAY49553:
XX
DT 13-JAN-2000 (first entry)
XX
DE Human endothelial membrane glycoprotein IIIa protein sequence.
XX
KW Human: coding sequence polymorphism; vascular pathology gene;
KW polymorphic site; phenotype correlation; forensic; paternity testing;
KW medicine; genetic analysis; vascular disease.
XX
XX Homo sapiens.
XX OS
XX PN W09950454-A2.
XX
XX 07-OCT-1999.
XX PD
XX PE 26-MAR-1999; 99WO-US06473.
XX PR 01-APR-1998; 98US-0054272.
XX
XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Lander ES, Daley CO, Cargill M, Ireland JS, Rozen SG;
XX WPI: 1999-620066/53.
XX DR N-PSDB; AAZ32162.
XX
XX Determination of polymorphisms in genes, especially those identifying
XX predisposition to vascular disease

PS Disclosure: Fig 6; 134pp; English.

CC AA232159 to AA232194 represent reference alleles for specifically
CC claimed nucleic acid sequences from the present invention which comprise
CC polymorphic sites as given in a table in the specification, selected
CC from 92 single nucleotide polymorphisms in which the nucleotide at the
CC polymorphic site is different from a nucleotide at the same site in a
CC reference allele. The nucleic acids, and primers and probes, are used to
CC identify polymorphisms, which may predispose an individual to disease,
CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AA49550 to AA49573 represent the proteins which correspond
CC to some of the reference alleles.

XX Sequence 788 AA:

Query Match 99.2%; Score 4119; DB 20; Length 788;
Best Local Similarity 99.3%; Pred. No. 6.5e-283;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCOOCLAVSPMCAMCSDEALPLGSPRODLKENLKNCAPESTIEPVS 60
DB 27 gpnicttrgyssccqclavspmcawcsdealpigsprcdlkenllkncapesiefvse 86
QY 61 ARVLEDRPLSDKSGSSOVTQVSPORIALRLRPDDSKNFSIOVROYEDYPVDIYYLMDL 120
DB 87 arvledrplsdksqsgssvqtvspqrlalrlrpdsknfsiqvqvdyvdiyyllmdl 146
QY 121 SYGMKDDLWSTONLGTAKTATQMKLTSNLRIGGAFVDKPVSPYMYISPEALENFCYDM 180
DB 147 sygmkdllwstlgnlgtaklqmkrltsnlrfgatvdkspspymyispealenpcydm 206
QY 181 KTTCLEPMFGYKHYVLTLDQVTRFNEEVKKOSVSRNRDAPEGGDAIMQATVCEKIGWRN 240
DB 207 ktcclepmfykhyvltldqvtrfneevkkqsvsrnrndapeggfdaimqatvcekiqwrn 266
QY 241 DASHLVEFTDAKTHIALDGRLAGIVOPNDGQCHVSDNHYSASTTMDYPSLGLMTEKLS 300
DB 267 dashllvftdaktahlaldgrlagivopndgqchvgsdnhyssasttmdypslglmekls 326
QY 301 OKNINILFANVENVMNLXONYSFLIPGTGVLSMDSSNVLIQIVAVGKIRSKYELVR 360
DB 327 oknlnlfavtenvmnllyqnyseflpghtvlyslmdssnvliqivdaykyltskvelevr 386
QY 361 DPEBELSLSPNATLNNENYVPGLSQMGKIGDVTFSFSLKAVRGCPQEKESFTIKPYG 420
DB 387 dpeelslslpnatlnnenyvpjlsqcmgkigdvtfsfslkavrgcpqekesftikpyg 446
QY 421 FKDSLIVQYTFDCDCACQAOAEPNSHRNCNNGTFEGGVCRGPGWLGSOCECSEEDYRP 480
DB 447 fkdslivqytfddccacqaaepnshrcnngtfeegvcrpgwlgsgcecseddyrp 506
QY 481 SOODESPREGPVCSORBECJCGQCVCHSSDPGRITGKRCDDPSCRYRGEMSGSG 540
DB 507 sqodecspregpvcsoorbecjcgqcvchssdpgritgkyrceddpscryrgemsgsg 566
QY 541 QCSGCDLSDSWTGYGTCCTRTDTCMSSNGLICSGRCKCEGSCVCTOPSPSYGTCK 600
DB 567 qcsgcdlsdswtgygctctrttdcmssnglicsgrckcegsctvctopsysygtck 626
QY 601 CPTCPACTFFKCEVCKKFFDRGALHDENTCNRYCRDEIESYKELKDTGKDAVNCTYKNE 660
DB 627 cptcpactffkceveckkffdrpymentcnrycrdeiesykelkdtgkdaavncitykne 686
QY 661 DDCVAFQYEEEDSSGSKILYVVEEPCPKGPRIIVLVLSYMAIILIGIAAILIKMLLT 720
DB 687 ddcvafiqyeeedssgskilyvveepecpkgpriivlvlsymailligiallikmlilt 746
QY 721 IHDRKEFAFEERARAKMDTANPLYKEATSTFTNITYRG 762
DB 747 ihdrkefakfeeararakmdtanplykeatstftnityrgt 788

RESULT 2

AAV49567
ID AAV49567 standard; Protein; 788 AA.

XX AAV49567;

XX 13-JAN-2000 (first entry)

XX Human antithrombin III protein sequence.

XX Human; coding sequence polymorphism; vascular pathology gene;

KW polymorphic site; phenotype correlation; forensic; paternity testing;

XX medicine; genetic analysis; vascular disease.

XX Homo sapiens.

XX WO9950454-A2.

XX 26-MAR-1999; 99WO-US06473.

XX 01-APR-1998; 98US-0054272.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX Lander ES, Daley CG, Cargill M, Ireland JS, Rozen SG;

XX WPL: 1999-620066/53.

XX N-PSDB: AA232186.

XX Determination of polymorphisms in genes, especially those identifying

XX predisposition to vascular disease

XX Disclosure: Fig 30; 134pp; English.

CC AA232159 to AA232194 represent reference alleles for specifically
CC claimed nucleic acid sequences from the present invention which comprise
CC polymorphic sites as given in a table in the specification, selected
CC from 92 single nucleotide polymorphisms in which the nucleotide at the
CC polymorphic site is different from a nucleotide at the same site in a
CC reference allele. The nucleic acids, and primers and probes, are used to
CC identify polymorphisms, which may predispose an individual to disease,
CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AA49550 to AA49573 represent the proteins which correspond
CC to some of the reference alleles.

XX Sequence 788 AA:

Query Match 99.2%; Score 4119; DB 20; Length 788;
Best Local Similarity 99.3%; Pred. No. 6.5e-283;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCOOCLAVSPMCAMCSDEALPLGSPRODLKENLKNCAPESTIEPVS 60
DB 27 gpnicttrgyssccqclavspmcawcsdealpigsprcdlkenllkncapesiefvse 86
QY 61 ARVLEDRPLSDKSGSSOVTQVSPORIALRLRPDDSKNFSIOVROYEDYPVDIYYLMDL 120
DB 87 arvledrplsdksqsgssvqtvspqrlalrlrpdsknfsiqvqvdyvdiyyllmdl 146
QY 121 SYGMKDDLWSTONLGTAKTATQMKLTSNLRIGGAFVDKPVSPYMYISPEALENFCYDM 180
DB 147 sygmkdllwstlgnlgtaklqmkrltsnlrfgatvdkspspymyispealenpcydm 206
QY 181 KTTCLEPMFGYKHYVLTLDQVTRFNEEVKKOSVSRNRDAPEGGDAIMQATVCEKIGWRN 240
DB 267 ktcclepmfykhyvltldqvtrfneevkkqsvsrnrndapeggfdaimqatvcekiqwrn 266
QY 241 DASHLVEFTDAKTHIALDGRLAGIVOPNDGQCHVSDNHYSASTTMDYPSLGLMTEKLS 300

```

DB 267 dashllvfttdakthialdgrlagivqpdngchvgsdnhysastctmdypslgimtekls 326
OY 301 OKNINLIFAVTENNVNLYQNTSELIPTGTVGLSMDSNVLIQI..VDAYGKIRKVELEVR 360
DB 327 qkninliffavtenvvnlyqnyseilipgtltvylsmdsenvliq..vdaygkirkvelevr 386
OY 361 DLPEELSLSFNATCLNNVIRPGKSCMGLKIGDVTFSFISIAKVRCCPOEKEKSTFIKRVG 420
DB 387 dlpeelslsfnatclnnevirpglkscmglkigdvtfsfisleakvrgcpqekesfikrvpg 446
OY 421 FKDSLIVOTFPCDCACQAQAPNSHRCNNGNGTEFCGVCRCGPGWLSQCCESEEDRP 480
DB 447 fkdslivgtfddcaqagapnshrcnngntfecgvcrcpgwlsqcceseedryp 506
OY 481 SQODECSPREGQPVCSQRCGLCGQCVCHSSDFKRTKTCYCCDDPFCVRYKGMCSGHC 540
DB 507 sqgdecspreqgpcvcsqrgcelcgcvchssdfgkltgkyccddfcsvrykgmcsghg 566
OY 541 QCSGDCCLCDSDMTGYVNCCTRTDTCMSSNGLCSGRKCEGSCVCIQPSYCDTCCK 600
DB 567 qscgdcclcdsdmtgyvncctrtctdcmssngllcsgrkcegsccvciqpsygdteck 626
OY 601 CPTCPDCTFKKCEVCKKFDGALHDENTCNRYCRDEIESYKELKDTGKDAVNCTYKNE 660
DB 627 cptcpdactfkkecvckkfdreymtentcnrycrdeiesvkelkdtgkdaavnclykne 686
OY 661 DDCVVFQYEDSSGKSTLYVEEPECPKGPDLVVLVSMVALILIGLALILYKLIT 720
DB 687 ddcvvvfqyedsdsgkstlyveepecpkygpdllvllsvmaaililglalilkyllit 746
OY 721 IHDRKEFAKFEERARAKWDTANPLYEATSTFNITYRGT 762
DB 747 ihdrkefakfeerarakwdtanpnykeatsfnityrgt 788

RESULT 3
AAV92443
ID AAV92443 standard; Protein: 788 AA.
AC AAV92443;
XX
XX 10-AUG-2000 (first entry)
DE Wild type human GPIIb, alloantigen PLAI.
XX
XX GPIIb: neurological disease; genotype: phenotype; diagnosis:
XX allantoigen; platelet antigen; PLAI; PLAI2: Alzheimer's disease; stroke;
XX neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
XX amyotrophic lateral sclerosis; multiple sclerosis; dementia.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX Peptide 1..26
XX Protein /Label= signal_peptide
XX /Label= mature_protein
XX
XX MO200020634-A1.
XX
XX 13-APR-2000.
XX
XX 01-OCT-1999; 99WO-1B01696.
XX
XX 01-OCT-1998; 98US-0102624.
XX
XX (NOVA-) NOVA MOLECULAR INC.
XX
XX Schapert K;
XX
XX WPI: 2000-303801/26.
XX
XX N-PSDB; AAA09362.

```

```

XX Identifying a subject at risk for a neurological disease comprises
PT determination of genotype or phenotype of GPIIb or GPIIb locus and
PT determining presence of variant GPIIb or GPIIb allele or isoform
XX
XX Disclosure; Fig 3; 55pp; English.
PS
XX
XX Two different forms of GPIIb, alloantigens PLAI and PLAI2 (for platelet
CC Antigen 1 and 2) have been described and can be distinguished using a
CC monoclonal antibody. The rarer form, PLAI2 (see AA09363), has sustained
CC a point mutation at base 192 that causes a nucleotide change from a "T" to
CC a C and thus a leucine to proline amino acid substitution at residue
CC position 33.
CC Identifying a subject at risk for a neurological disease comprises
CC determining the genotype or phenotype of the GPIIb or GPIIb locus of a
CC subject, and determining the presence of a variant GPIIb or GPIIb
CC allele or isoform, where the presence indicates an increased risk of
CC neurological disease. The method is useful for identifying a risk,
CC diagnosing and treating a neurological disease comprising Alzheimer's
CC disease, neurofibromatosis, Huntington's disease, depression,
CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
CC disease and multi-infarct dementia.
XX
XX Sequence 788 AA:
SQ
Query Match 99.24; Score 4119; DB 21; Length 788;
Best Local Similarity 99.34; Pred. No. 6,5e-283;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 GPNICTTRGVSSCOOCLAVSMCMWCSDEALPLCSFPCDLKENLKNCAPESTIEFVSE 60
DB 27 gpnicttrgvssccqlavsmcmwcsdealpysprcdlkenllkncapeslefpvse 86
OY 61 ARVLEDRPLSDKSGDSQVQVSPORTALRLRPDDSKNFESIUVQVEDYVVDIYVIMDL 120
DB 87 arvledrplsdksqdsqvvqvsportalrlrpdsknfsiqvqvedyvdiylmldl 146
OY 121 SYSMKDLMSIQNLGKTLATQMRKLTSLNRIGFAGFVDKPVSPYMYISPPALENPCYDM 180
DB 147 sysmkddlswisqnlgtklatqmrkltslnrltfgafgfdkpvspymyisppaleenpcydm 206
OY 181 KTTCLPMFGYKHVLTLDQVTRFNEBVKKQSVSRNRDAPBEGUAIKATYCDCKIGWRN 240
DB 207 ktclpmfgykhvltldqvtrfneebvkkqsvsrnrdapeguailkatycdckigwrn 266
OY 241 DASHLIVTTDAKTHIALDGRLAGIVQPDNGCHVGSNDNRASATMDYPSLIGMTFKLS 300
DB 267 dashllvtttakthialdgrlagivqpdngchvgsdnhysastctmdypslgimtekls 326
OY 301 OKNINLIFAVTENNVNLYQNTSELIPTGTVGLSMDSNVLIQI..VDAYGKIRKVELEVR 360
DB 327 qkninliffavtenvvnlyqnyseilipgtltvylsmdsenvliq..vdaygkirkvelevr 386
OY 361 DLPEELSLSFNATCLNNVIRPGKSCMGLKIGDVTFSFISIAKVRCCPOEKEKSTFIKRVG 420
DB 387 dlpeelslsfnatclnnevirpglkscmglkigdvtfsfisleakvrgcpqekesfikrvpg 446
OY 421 FKDSLIVOTFPCDCACQAQAPNSHRCNNGNGTEFCGVCRCGPGWLSQCCESEEDRP 480
DB 447 fkdslivgtfddcaqagapnshrcnngntfecgvcrcpgwlsqcceseedryp 506
OY 481 SQODECSPREGQPVCSQRCGLCGQCVCHSSDFKRTKTCYCCDDPFCVRYKGMCSGHC 540
DB 507 sqgdecspreqgpcvcsqrgcelcgcvchssdfgkltgkyccddfcsvrykgmcsghg 566
OY 541 QCSGDCCLCDSDMTGYVNCCTRTDTCMSSNGLCSGRKCEGSCVCIQPSYCDTCCK 600
DB 567 qscgdcclcdsdmtgyvncctrtctdcmssngllcsgrkcegsccvciqpsygdteck 626
OY 601 CPTCPDCTFKKCEVCKKFDGALHDENTCNRYCRDEIESYKELKDTGKDAVNCTYKNE 660
DB 627 cptcpdactfkkecvckkfdreymtentcnrycrdeiesvkelkdtgkdaavnclykne 686

```

QY 661 DDCVRFQYEDSSGSKSLVVEEPCPKGPDILVLLSVGAILLGLAALLIMKLLIT 720
 Db 687 ddcvrrfygddsgsksllyvveeepckpddilvllsvmgallllglaallimkllit 746
 QY 721 IHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNTTYRG 762
 Db 747 ihdrkefakfeeararakwdtannplykeatsftnttyrgt 788
 RESULT 4
 ID AAY81461 standard; Protein; 788 AA.
 XX AAY81461;
 AC 03-JUL-2000 (first entry)
 DE Human integrin beta 3.
 DE Integrin beta 3; human endothelial glycoprotein; GP3A; GPIIb;
 KM ITGB3; CD61; platelet glycoprotein 3a; cellular adhesion;
 KM vitronectin receptor; fibronectin receptor; expression inhibition;
 KM antisense therapy; tumour formation; cancer invasion; bleeding disorder;
 KM inflammation.
 XX Homo sapiens.
 OS
 XX US6037116-A.
 PN 14-MAR-2000.
 PD 25-JUN-1999; 99US-0344520.
 PF 25-JUN-1999; 99US-0344520.
 PR (ISIS-) ISIS PHARM INC.
 XX (ISIS-) ISIS PHARM INC.
 PA Bennett CF, Cowsett LM, Monia BP;
 PI WPI: 2000-246189/21.
 DR N-PsDB; AAA07028.
 PT New antisense compound that inhibits human integrin beta3, useful e.g.
 PT for treating or preventing infection, inflammation and tumors
 XX
 PS Disclosure; Columns 43-48; 33pp; English.
 CC This sequence represents human integrin beta 3. Integrins constitute
 CC one of four classes of cellular adhesion molecules, and play an
 CC important role in cell migration, cell anchorage to substrates and
 CC cytoadhesion signalling pathways. They are heterodimeric
 CC cation-dependent membrane glycoproteins composed of an alpha and beta
 CC subunit. Integrin beta 3 (also known as human endothelial glycoprotein,
 CC GP3A, GPIIb, ITGB3, CD61 and platelet glycoprotein 3a) is the common
 CC beta subunit partner of the members of the beta-3 subfamily of integrins.
 CC This family consists of the vitronectin receptor (alpha-V-beta-3) and the
 CC fibronectin receptor (alpha-IIb-beta-3). Cells expressing this class of
 CC integrin can adhere to various matrix proteins and participate in
 CC various cytoadhesion-driven cellular responses. Integrin beta 3 is
 CC implicated in conditions such as vascular restenosis, excessive bone
 CC resorption, angiogenesis (in melanoma), tumour invasion, platelet
 CC aggregation and glanzmann's thrombasthenia. The invention relates to
 CC antisense oligonucleotides targeted to the human integrin beta 3 gene,
 CC which inhibit its expression. A series of oligonucleotides
 CC (AAA07035-A07074) were designed to target different regions of the human
 CC integrin beta 3 RNA, and were analysed for their effect on integrin beta
 CC 3 mRNA levels by quantitative real-time PCR. The oligonucleotides of the
 CC invention are useful for diagnosis, prevention and treatment of
 CC conditions associated with integrin beta 3 expression, such as tumour
 CC formation, inflammation, infections and the diseases mentioned above.
 CC
 SQ Sequence 788 AA;

Query Match 99.2%; Score 4119; DB 21; Length 788;
 Best Local Similarity 99.3%; Fred. No. 6.5e-285;
 Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 GPNICTTRGVSSCOOCLAVSPMCAMCSDEALPLGSPKCDLKENLMDKCAPESIEFPVSE 60
 Db 27 gpnicttrgvsscgqclavspmcawcsdealpigsprcdlkenlmdkcapesiefpvse 86
 QY 61 ARVLEDRLSDKSGSDSSQVTVSPQRIARLRPDDSKNF51 QVRQVEDYVVDIYIMDL 120
 Db 87 arvledrplsdksqsdssqvtvpsqrialrlrpdsknf51qvrqvedyvvdilyimdl 146
 QY 121 SYSMKDDLSIQNLGTAKATQKRKTSNLRIGFGAFVYKPKVPYKPIKYSPPALENPGCDM 180
 Db 147 sysmkddlswsiqnlgtakatkrltsnlrigrfgafvykpkvpypkpiyksppalenpgcdm 206
 QY 181 KTTCLPMFGYKHHVLTLPQVTFNFEVKKOSVSRNRRLAPEGGUAIMQATVCDERIGWRN 240
 Db 207 ktclpmtfgykhhvltlpqvtfneevkksvsnrnrldapeggidaimgatvcdetigrwn 266
 QY 241 DASHLLVFTTDAKTHIALDGLAGIVQPNDSQCHVGSNNHYSASTMDYPSLGIMTEKLS 300
 Db 267 dashllvfttdakthialdglagivqpnmsgchvgsnnhysastmdypslgimtekls 326
 QY 301 OKNINLIRAVTENYVNLQONSELIPQTVGLSMDSSNNVQLIYDAVGRKRSVELEVR 360
 Db 327 oknlnliravtenvnylnqonseilpqtvglsmssnnvqllydavgkrksrvelevr 386
 QY 361 DLPEELSLSPNATCLNNEVIRIGLSCMGLKIGDVTNFIIEKAVNGCPQEKRSSTIRPVG 420
 Db 387 dlpeelstspnatclnnevirlglsclmglkigdvtnfieekavngcpqekrsstirpvg 446
 QY 421 FKDSLIVVTPDDCACQAQAEAPNSHRCNNGNPFEGCVRCRCRGTWLSQCESEEDYRP 480
 Db 447 fkdslivvtpddcacqaqaepnshrcnngnpgfegcvrcrcrgtwlsgqceceedyrrp 506
 QY 481 SQODECSPREGOPVCSOGECICGOCVCHSSDPFKITGKVCDFGFCVRRKGMCSGHG 540
 Db 507 sqodecspregopvcsogecicgocvchssdpfkrtgkvcdfgfcvrrkgmcsghg 566
 QY 541 QCSGDCLCDSDWMTGYCNCRTTRDTGCMSSNGLSCGKCEGCVSVCIOGSGYDTCER 600
 Db 567 qcsgcdclcdsdwmtgycncrttrdtgcmssnslscgkcegcvsvcioqsgydtcer 626
 QY 601 CPTCPDACTFRKCEVECKEPDRGALHDETCNRCRDEFSVRLKDTGDAVNTCYKNE 660
 Db 627 cptcpdactfrkceveckepdrgalhdeTCNRCRDEFSVRLKDTGDAVNTCYKNE 686
 QY 661 DDCVRFQYEDSSGSKSLVVEEPCPKGPDILVLLSVGAILLGLAALLIMKLLIT 720
 Db 687 ddcvrrfygddsgsksllyvveeepckpddilvllsvmgallllglaallimkllit 746
 QY 721 IHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNTTYRG 762
 Db 747 ihdrkefakfeeararakwdtannplykeatsftnttyrgt 788
 RESULT 5
 ID AAY92451 standard; Protein; 762 AA.
 XX AAY92451;
 AC 10-AUG-2000 (first entry)
 DE GPIIb variant Arg214Gln.
 DE GPIIb variant Arg214Gln.
 KW GPIIb; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KW allantoigen; platelet antigen; PLAI; PLAI2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
 KW anyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.
XX
XX W0200020634-AL.
XX
PD 13-APR-2000.
XX
PE 01-OCT-1999; 99MO-IB01696.
XX
PR 01-OCT-1998; 98US-0102624.
XX
PA (NOVA-) NOVA MOLECULAR INC.
XX
PI Schappert K;
DR WPI; 2000-303801/26.
XX
PT Identifying a subject at risk for a neurological disease comprises
XX determination of genotype or phenotype of GPIIIA or GPIIb locus and
XX determining presence of variant GPIIIA or GPIIb allele or isoform

Disclosure; Page -: 55pp; English.

Known polymorphisms in GPIIIA that may be determined to be variants
using the methods of the invention include AAY92447-56, which are variant
GPIIIA sequences derived from the mature protein.
Identifying a subject at risk for a neurological disease comprises
determining the genotype or phenotype of the GPIIIA or GPIIb locus of a
subject, and determining the presence of a variant GPIIIA or GPIIb
allele or isoform, where the presence indicates an increased risk of
diagnosing and treating a neurological disease comprising Alzheimer's
disease, neurofibromatosis, Huntington's disease, depression,
amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
disease and multi-infarct dementia.
Note: This sequence is not given in the specification it was created
from the wild type human GPIIIA sequence which appears in Figure 3.

Sequence 762 AA;

Query Match	99.1%	Score 4115:	DB 21:	Length 762:
Best Local Similarity	99.2%	Pred. No. 1.2e-282:		
Matches 756:	Conservative 1:	Mismatches 5:	Indels 0:	Gaps 0
Qy	1	GNICTTTRGSSCOOC	LAVSPMCAMCSDEAL	LGSPRCDLKENLKDNCAPSEIEPVSE 60
Db	1	gnictctgsvascqgc	lavsmpmcawcsdeal	lpgsprecldkenllkdnccapsieiefpvse 60
Qy	61	ARVLEDRPLSDKGS	SSQVTVQVSRQRLRLRPDS	KNFSTQVRVEDYRPDIYYLML 120
Db	61	arvledrplsdqsgds	sqvvtqvsrpqrlrlrlpds	knfstsivrvedyvpdiyyimdl 120
Qy	121	SYSMKDDLSIONL	STKTLKTQMRKLTSLNLRIG	FAFVDRPVPSPYMTSPPEALENCYCM 180
Db	121	ysmmdldltsiqn	lgtklktatqmrltslnrl	lfgatrvckpvsppymylsppealencpccym 180
Qy	181	KTTCLPMEGYKAV	LLTDOVTRPNEEYKAKOS	SRNRPDAEGGFDAIMQATVCDEKIGMRN 240
Db	181	kttclpmyfkyhnl	tlctdvttrfneeevkkqvsq	nrdapeggfdaimgatvcdcklqwtcn 240
Qy	241	DASHLLEVTDAKTH	IALDGRLAGIYQPNDDGCH	SVSDNHYSASTMDDVPSLCILMEKTS 300
Db	241	dashllvftcdketh	ialdgrltaglvgpnddgchv	srdnhysasttmddvpslglmekts 300
Qy	301	QKNINLIEAVTEN	VNLYONYSGLIPGTTVGL	LSMDSSNVLOLIDAYGKIRNSVELEVR 360
Db	301	qkninllfavten	vnvnllyngysellpgtct	vglismdsnvlqlivdaygkirtsvelevr 360
Qy	361	DLPEELISFNMT	CLANNEYTPGLKSCMGKIG	TVFSTFEAKVRGQPOKEKEKFTTKPVG 420
Db	361	dlpeelistsfnmt	clannevtpglkscmgkig	tvfstsfeakvrgqpokekfstklkpv 420

Oy		421 FKSDLIYQVWFEDCDCAOAOAEFNSHRCNNGNTFCGVCRCGPMLISOCSCSEDRP	480
Dd		421 fKsdLiYqvLfEdcdCaCqaqAdEpshrcmngntfEcgvCrGpYwIsqcCeSeDrP	480
Oy		481 SQDECCSPREGOPVCSQRGBECLCGQCCHSHDFGKITGKYCECDPFSCVRRKGEMCSHG	540
Dd		481 sqgdeCsprEqgpvcagrgscLsgcgvchssdftgkltgyJeceddIfScvrykgemcsgh	540
Oy		541 QCSCGDCLDSDMWTGYCCNCTTPTDTCMSSNGLLCSGRKCCGCCGCVCITPGSYGDTCEK	600
Dd		541 qcscgdclcdsdwtgyccnctttrdtcmssngllcsgrgkcgcgsccvcitppsygdteck	600
Oy		601 CPTCPDACTKKKEVECKKFRDGRALHDENMCNRYCDELSEVKELKDIDGKAIVNCTYNE	660
Dd		601 cPtcpdactfkkeveckktfdrepmhtlcnrYcedlesvkElkdIdgkdAvncTyne	660
Oy		661 DDCVRFQYVEDSSGSKILVLVEEPCCPKGPDIILVLLSVMGAILLICHAALLIKWLIT	720
Dd		661 ddcvrrfygyedssgskilylvveepccpkgpdiilvvllsvmgaillichaallikwlilit	720
Oy		721 IHDRREFAKFEERARAKWDTPANNPLYKEATSTFTNIITYRGT	762
Dd		721 ihdrkefakfeeerarakwdtpannplykeatsftftniityrgt	762
RESULT	6		
ID	AARS1500		
XX	AARS1500 standard; Protein; 788 AA.		
AC	AARS1500;		
DT	17-NOV-1994 (first entry)		
XX			
DE	Human platelet GPIIb surface antigen Yuk-a.		
XX			
KM	Platelet surface antigen: glycoprotein GPIIb; Yuk antigen; probe:		
KM	detection: thrombocyte, Yuk-a allele.		
XX			
OS	Homo sapiens.		
XX			
FH	Key Location/Qualifiers		
FT	Misc-difference 169 /note= "Yuk-a and Yuk-b differ at position 169 with		
FT	Gln (corresp. to CNA codon) in Yuk-a but		
FT	Arg (corresp. to CGA codon) in Yuk-b"		
XX			
PN	JP06078771-A.		
XX			
PD	22-MAR-1994.		
XX			
PF	06-JUL-1991; 91JP-0192554.		
XX			
PR	06-JUL-1991; 91JP-0192554.		
XX			
PA	(JUUI/) JUUT T.		
XX			
DR	WPI; 1994-131276/16.		
XX	N-PSDB; AAO62199.		
PT	The detection of platelet surface antigen determinant - using		
PT	oligo-nucleotide probes for Yuk a and Yuk b		
PS	Disclosure; Fig 1-5 (Page 8-12); 12pp; Japanese.		
XX			
CC	Primers (3) and (4) (AAO62195 and AAO62196, respectively) were used to		
CC	amplify a region of the yuk gene from platelet cDNA derived from		
CC	humans of known yuk antigen type. The 1mer probes AAO62193 and		
CC	AAO62194, specific for the yuk-a and yuk-b alleles respectively, were		
CC	tested for hybridisation to the cDNA. The yuk-a probe hybridised to		
CC	cDNA from a yuk a/a homozygote and from yuk a/b heterozygotes only,		
CC	while the yuk-b probe hybridised to cDNA from the heterozygotes and		
CC	from yuk b/b homozygotes only. AARS1500 is the sequence for the human		
CC	yuk-a protein.		

XX Sequence 788 AA;

Query Match 99.1%; Score 4115; DB 15; Length 788;
Best Local Similarity 99.2%; Pred. No. 1.2e-282;
Matches 756; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPIGSPKDLKENILKDNCAPISEIPEVSE 60
DB 27 gpnictrgvsscgqclavspmcawcsdealpigsprcdlkenllkdncapestiefpvse 86
QY 61 ARVLEDRPLSDKSGSSQVTOVSFORIALRLRPDSSKNFSIQVRQVEDYPVDIYYLMDL 120
DB 87 arvledrplsdksqsgssqvtqvsprlrlrldpssknfsiqvrqvedyvdiyylmldl 146
QY 121 SYSMKDLMSIQNLGFKLATOMRKLTSLNRIGFAGVDFVKPVPYMYTISPPEALENPCYDM 180
DB 147 symskdllswisqnlglklatqmkltslnrifgafvdfvpvpymytisppealenpcydm 206
QY 181 KTTCLPMEFGYKHVLLTDQVTRFNEEVKKOSYRNNDAPEGGFDAIMQATVCEKIGWEN 240
DB 207 ktclpmefgykhvlltdqvtfrneevkkosyrnndapeggfdaimqatvcekiqwn 266
QY 241 DASHLLVFTTDAKTHIALDGRLAGIYQPNDSQCHVSDNHYSASTTMDYPSLGLMTEKLS 300
DB 267 dashllvfttdakthialdgrlagiypndsgchvsnhysasttmdypsigtmtkls 326
QY 301 OKNINLIFAVTEWVNLVONYSELIPGTVGVLSMSSWVLIOLIVDAVGKIRSKYLEVR 360
DB 327 qkninlifavtenvnlvnyseellpgtlvgvlsmdsswvliolvdaygkirkylevr 386
QY 361 DLPEELSLSFNATCLNNEVYIPGLKSCMGLIKIDTVSFSIEAKVRGCPQEKESFTIKPVG 420
DB 387 dlpeelslsfnatclnnevypglksomglkigtvsfsieakvrqcpqekesftikpv 446
QY 421 FDSLIIYQVTFPCDDACAQAAPNHRNNGNTEGCVRCRCGPGWLGSGQCESEEDYRP 480
DB 447 fksliivqvtlfcddacagaaepnshrcnngntecgvcrcgpgwlgsgceceedyrp 506
QY 481 SQODECSPREGQPVCSORGECLGQCVCHSSDFGKITKGCPCDDFSCVRRKKGEMCSHG 540
DB 507 sqodecspregqpvcsqgeclgqcvchssdfgkitkgyccddfscvrrkkgemcsghg 566
QY 541 QCSGCDLCLDSDMWTGYNCNCTRTTDCMSSNMLCSGRKCEGSCVCIQPGSYGDTTEK 600
DB 567 qcsgcdclcdsdmwtgyncnctrttcdmssnmlcsgrkcegcscvcicpgsygdttek 626
QY 601 CPTGCDACTFKKCEVCKKFDRLGALHDENTCNRICRDEIESVVKELKDKDAVNCYKNE 660
DB 627 cptgcdactfkkecvckkfdregymtenctnrycrdeiesvkelkdkdavnctykne 686
QY 661 DDCVVRFOYEDSSGKSLIYVEEPECKPDPDILVLLSWGAILLGLALLIMKLLIT 720
DB 687 ddcvvrffoyedsqsksllyveepecpdpdillvllswgaillglallimkllit 746
QY 721 IHRKREFAKFEERARAKMDTANNPLYKEATSTFTNTITRYGT 762
DB 747 ihdrkefakfeerarakmdtanplykeatstftntitrygt 788

```

RESULT 7
AAV92456
ID AAV92456 standard; Protein: 762 AA.

AC AAV92456;

DT 10-AUG-2000 (first entry)

XX GPIIa variant Ser752Pro.

KW GPIIa: GPIIb: neurological disease; genotype: phenotype; diagnosis:
alloantigen; platelet antigen; FLAI; FLA2; Alzheimers disease; stroke;

KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.

PN WO200020634-A1.

PD 13-APR-2000.

PE 01-OCT-1999; 99WO-IB01696.

PR 01-OCT-1998; 98US-0102624.

PA (NOVA-) NOVA MOLECULAR INC.

PI Schappert K;

DR WPI: 2000-303801/26.

PT Identifying a subject at risk for a neurological disease comprises

PT determination of genotype or phenotype of GPIIIa or GPIIb locus and

XX determining presence of variant GPIIIa or GPIIb allele or isoform

XX Disclosure; Page -: 55pp; English.

XX Known polymorphisms in GPIIIa that may be determined to be variants

XX using the methods of the invention include AAV92447-56, which are variant

XX GPIIIa sequences derived from the mature protein.

XX Identifying a subject at risk for a neurological disease comprises

XX determining the genotype or phenotype of the GPIIIa or GPIIb locus of a

XX subject, and determining the presence of a variant GPIIIa or GPIIb

XX allele or isoform, where the presence indicates an increased risk of

XX neurological disease. The method is useful for identifying a risk,

XX diagnosing and treating a neurological disease comprising Alzheimer's

XX disease, neurofibromatosis, Huntingdon's disease, depression,

XX amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's

XX disease and multi-infarct dementia.

XX Note: This sequence is not given in the specification it was created

XX from the wild type human GPIIIa sequence which appears in Figure 3.

SQ Sequence 762 AA;

Query Match 99.0%; Score 4114; DB 21; Length 762;
Best Local Similarity 99.2%; Pred. No. 1.4e-282;

Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPIGSPKDLKENILKDNCAPISEIPEVSE 60
DB 1 gpnictrgvsscgqclavspmcawcsdealpigsprcdlkenllkdncapestiefpvse 60
QY 61 ARVLEDRPLSDKSGSSQVTOVSFORIALRLRPDSSKNFSIQVRQVEDYPVDIYYLMDL 120
DB 61 arvledrplsdksqsgssqvtqvsprlrlrldpssknfsiqvrqvedyvdiyylmldl 120
QY 121 SYSMKDLMSIQNLGFKLATOMRKLTSLNRIGFAGVDFVKPVPYMYTISPPEALENPCYDM 180
DB 121 symskdllswisqnlglklatqmkltslnrifgafvdfvpvpymytisppealenpcydm 180
QY 181 KTTCLPMEFGYKHVLLTDQVTRFNEEVKKOSYRNNDAPEGGFDAIMQATVCEKIGWEN 240
DB 181 ktclpmefgykhvlltdqvtfrneevkkosyrnndapeggfdaimqatvcekiqwn 240
QY 241 DASHLLVFTTDAKTHIALDGRLAGIYQPNDSQCHVSDNHYSASTTMDYPSLGLMTEKLS 300
DB 241 dashllvfttdakthialdgrlagiypndsgchvsnhysasttmdypsigtmtkls 300
QY 301 OKNINLIFAVTEWVNLVONYSELIPGTVGVLSMSSWVLIOLIVDAVGKIRSKYLEVR 360
DB 301 qkninlifavtenvnlvnyseellpgtlvgvlsmdsswvliolvdaygkirkylevr 360
QY 361 DLPEELSLSFNATCLNNEVYIPGLKSCMGLIKIDTVSFSIEAKVRGCPQEKESFTIKPVG 420

```

Db 361 dlpelalsfnatclnnevlpglkscmglkigtvtsfisleakvrgcpqekesftlkpyvg 420
 Oy 421 FKDSLIVQVTFDDCACAOAEPNSHRCNNGNGFEGCVRCRCGPMGSGOCESEEDYRP 480
 Db 421 fkdslivqvtfdcdcaqaqaepnshrcnmngntfegvrcgpgvvgsgceceedytr 480
 Oy 481 SQODECSPREGQPVCSQRCGLCGQCVCHSSDFGKITGKYCECDPFCVRYKGMESGHC 540
 Db 481 sqgdecspregqpvcsqrgcelcgcvchssdfgkitgkycecdpfcvrykgemcsghg 540
 Oy 541 QCSGDCCLCDSDMWTGYCNCCTTRDTCMSSNGLLCSGRGKCEGSCVCYTOPGSYGTCEK 600
 Db 541 qcsgcdclcdsdwtygnccttrdtcmssngllcsgrgkcegsccvcytopgsygtceek 600
 Oy 601 CPTCDACCTFKKECECKKFRGALHDENTCNRCRDIEISVKEKLDKTDKDAVNCYKNE 660
 Db 601 cptcdactfkkececkkfdreymtencnrycdelesvkelktdgkdaavncykne 660
 Oy 661 DDCVVRFOYEDSSGKSILYVEEPECPKGPDLVLLSVMGAILLIGLAALLIMKLIT 720
 Db 661 ddcvvrfoyyedssgksillyveepecpkgpdlvllsvmgailllglaallimkllit 720
 Oy 721 IHDRKEFAKFEERARAKWDTANNPLLYKEATSTFTNTTYRGT 762
 Db 721 ihdrkefakfeerarakwtdannplykeatstftntlyrgt 762

RESULT 8

AAV92448
 ID AAV92448 standard; Protein: 762 AA.

AAV92448;

DT 10-AUG-2000 (first entry)

DE GPIIa variant Leu117Trp.

XX GPIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;

KW allonigen; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;

KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;

KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.

PN WO200020634-A1.

PD 13-APR-2000.

PF 01-OCT-1999; 99WO-1B01696.

PR 01-OCT-1998; 98US-0102624.

PA (NOVA-) NOVA MOLECULAR INC.

PI Schappert K;

DR WPI: 2000-303801/26.

XX Identifying a subject at risk for a neurological disease comprises

PT determination of genotype or phenotype of GPIIa or GPIIb locus and

PI determining presence of variant GPIIa or GPIIb allele or isoform

XX Disclosure: Page -: 55pp; English.

CC Known polymorphisms in GPIIa that may be determined to be variants

CC using the methods of the invention include AAV92447-56, which are variant

CC GPIIa sequences derived from the mature protein.

CC Identifying a subject at risk for a neurological disease comprises

CC determining the genotype or phenotype of the GPIIa or GPIIb locus of a

CC subject, and determining the presence of a variant GPIIa or GPIIb

CC allele or isoform, where the presence indicates an increased risk of

CC neurological disease. The method is useful for identifying a risk,

CC diagnosing and treating a neurological disease comprising Alzheimer's

CC disease, neurofibromatosis, Huntingdon's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification. It was created
 CC from the wild type human GPIIa sequence which appears in Figure 3.

XX Sequence 762 AA:

Query Match 99.0%; Score 4113; DB 21; Length 762;

Best Local Similarity 99.2%; Pred. No. 1,66-282;

Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPGSPRCDEKENLDMNCAPEISIEPVE 60
 Db 1 gpnictrgvsscgqclavspmcawcsdealpysprcdeknlldmncapesiepvse 60
 Oy 61 ARVLEDRPLSDKSGSSQVTOYSPQRIARLPDSDKNSTIVROYEDYPVDIYLMDL 120
 Db 61 arvledrplsdksqssqvtqyspqrlarlppdsdknstsivroyedypvdiylmmdl 120
 Oy 121 SYSWKDLMSTIONLGTKLATQMRKITSNLRIFGAFVDRKPVSYMYTISPEALENFCYDM 180
 Db 121 syswkdlmstionlgtklatqmrkitsnlrifgafvdrkpvsymytispealenfcydm 180
 Oy 181 KTTCLPMFGYKHWLTLTDQVTRNEEYKQSVSRNRDAPEGHDAIMQATVCOEKIGWRN 240
 Db 181 ktclpmyfgykhwltltldqvtrneeykqsvsrnrdapeghdaimgatvcoekigwrn 240
 Oy 241 DASHLVFTTDDATHTALDGRLAGIYQPNDOCHVSDNHTSSTYMDVPSGLMTEKLS 300
 Db 241 dashlvfttdathtalldgrlagiyqpnodochvsdnhtsstymdvpsglmteklss 300
 Oy 301 QKNINLFAVTENVNLKYMYSELIPTGTGVLSMDSNVYQLVDAYGKTRSKVELEVR 360
 Db 301 qkninlfahtenvnlyqmyselipgtgvlsmdsnvylqivdaygktrskvelevr 360
 Oy 361 DLPELSLSFNATCLNNEVIRPLKSCGKLGKIGTVSEFSIEAKVRCQDEKESFTIKPYG 420
 Db 361 dlpelalsfnatclnnevlrplkscmgllkigtvtsfisleakvrgcpqekesftlkpyg 420
 Oy 421 FKDSLIVQVTFDDCACAOAEPNSHRCNNGNGFEGCVRCRCGPMGSGOCESEEDYRP 480
 Db 421 fkdslivqvtfdcdcaqaqaepnshrcnmngntfegvrcgpgvvgsgceceedytr 480
 Oy 481 SQODECSPREGQPVCSQRCGLCGQCVCHSSDFGKITGKYCECDPFCVRYKGMESGHC 540
 Db 481 sqgdecspregqpvcsqrgcelcgcvchssdfgkitgkycecdpfcvrykgemcsghg 540
 Oy 541 QCSGDCCLCDSDMWTGYCNCCTTRDTCMSSNGLLCSGRGKCEGSCVCYTOPGSYGTCEK 600
 Db 541 qcsgcdclcdsdwtygnccttrdtcmssngllcsgrgkcegsccvcytopgsygtceek 600
 Oy 601 CPTCDACCTFKKECECKKFRGALHDENTCNRCRDIEISVKEKLDKTDKDAVNCYKNE 660
 Db 601 cptcdactfkkececkkfdreymtencnrycdelesvkelktdgkdaavncykne 660
 Oy 661 DDCVVRFOYEDSSGKSILYVEEPECPKGPDLVLLSVMGAILLIGLAALLIMKLIT 720
 Db 661 ddcvvrfoyyedssgksillyveepecpkgpdlvllsvmgailllglaallimkllit 720
 Oy 721 IHDRKEFAKFEERARAKWDTANNPLLYKEATSTFTNTTYRGT 762
 Db 721 ihdrkefakfeerarakwtdannplykeatstftntlyrgt 762

RESULT 9

AAV92450
 ID AAV92450 standard; Protein: 762 AA.

XX AAV92450;

DT 10-AUG-2000 (first entry)

XX GP11a variant Ser162Leu.
 DE
 XX
 KW GP11a; GP11b; neurological disease; genotype; phenotype; diagnosis;
 KW allantoigen; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntingdon's disease; depression; Parkinson's disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 XX
 OS Homo sapiens.
 XX
 PN WO200020634-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-IB01696.
 XX
 PR 01-OCT-1998; 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K;
 DR WPI; 2000-303801/26.
 XX
 PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GP11a or GP11b locus and
 PT determining presence of variant GP11a or GP11b allele or isoform
 XX
 PS Disclosure; Page -: 55pp; English.
 XX
 CC Known polymorphisms in GP11a that may be determined to be variants
 CC using the methods of the invention include AAY92447-56, which are variant
 CC GP11a sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GP11a or GP11b locus of a
 CC subject, and determining the presence of a variant GP11a or GP11b
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntingdon's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GP11a sequence which appears in Figure 3.
 CC
 XX
 SQ Sequence 762 AA:

Query Match 99.0%; Score 4113; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 1.6e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICITRGVSSGQCLAVSPMCAMCSDEALPLGSPRDLKKNLKNKNCAPSIFFPVS 60
 DB 1 gp1ictlyvsscgcclavspmcawcsdealp1gsp1rcldkenllkdcncapeslefpyse 60
 QY 61 ARVLEDRPLSDKSGSDSSQVTSVSPORIALRLRPDDSKNFSIQVQVEDYPDVIYIMDL 120
 DB 61 arvledrplsdksdsgsdssqvtvsporlrlrppddsknfsiqvqvvedypvdiylmldl 120
 QY 121 SYSMKDLMSIONLGTIKTATQMRKLTSLNIGFAGFVDPKVPSPYWIISPPEALENPPCYDM 180
 DB 121 sysmkddlmsionlgtiktatqmrkltslnigfagfvdvpkvp1pywiisppealenpcydm 180
 QY 181 KTTCLPMFGYKHVLTITQVTRFNEEVKKQSSRRNDAPEGFDALMQATVDEKIGWRN 240
 DB 181 ktclcpmfgykhvltitqvtrfneevkkqssrrndapegfdalmaqatvdekigrwn 240
 QY 241 DASHLVFTTDAKTHIALDLGRLAGIYQPNDGCHVSDHNYASTTMDYPSGLTKEKIS 300
 DB 241 dshlvfttdakthialdlgrlagiyqpnodgchvshnyasttmdypsgltkeklis 300
 QY 301 QRNINLFAVTENVVNLQYNSYSELIPGTTGVLSMDSSNVQLIVDAYGKIRSKVELEVR 360

DB 301 qknunllfaveenvvnlyqnyse1iprtvgtvlsmddsnv1q1vdaygk1rskvelevr 360
 QY 361 DLPEELSLISFNATCLNNEVYIPGLKSCMGLKIGDVIYSFTEKAVNGCPEKSKSTIRPVG 420
 DB 361 dlpeelslisfnatclnnev1pglkscomglk1gdtvysfsteakvngcpkeksftl1rpvg 420
 QY 421 FRDSLIVQTPDCACAOAEPNSHRCNNMGFECGVCACGCGWGLSCECESEEDYRP 480
 DB 421 frdslivqtpdcacaaopnsrhrcnmgfegvcvccgpgw19sgceceedyrp 480
 QY 481 SQODECSPREGQPVCSQREGLCGVCVCHSSDFKITGKYCECDFFSCVRRYKEMCSGAG 540
 DB 481 sqodecspregegvpvcsqrgeclcgvcvchssdfgk1gkycecdffscvrykgemcsghg 540
 QY 541 QCSGCDCLDSDMWTGYTNCITTRDTCMSNGLLCSGKRGKCGSCYCIQPSIGDPTCEK 600
 DB 541 qcsgcdclcdsdmwtgytncittrdtcmsngllcsgrgkcgscycvciqpsygdptcek 600
 QY 601 CPTCPDACPFRKKECECKKFDKRGALHDENTCNRYCDEIESKFKLDPKDAVACTYKNE 660
 DB 601 cptcpdactfrkceveckkfdrepymtentcnrycdeiesvkelkdpkdaavactykne 660
 QY 661 DDCVVRFOYEDSSGKSLIYVEEPECPKGPDIIVLLSYGAILLGLAALLIWKLLIT 720
 DB 661 ddcvvrfoyedssgksliyyveeepcpkdpdiivllswygaill1glaall1wkllit 720
 QY 721 IHDKEFAKFEEDERAKAKWDTANPLYEKATSTNTNTYRGT 762
 DB 721 ihdrkefakfeederarakwtdanplykeastntntlyrgt 762

RESULT 10
 AAY92444
 ID AAY92444 standard; Protein; 788 AA.
 XX
 AC AAY92444;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Variant human GP11a, allantoigen PLA2.
 XX
 KW GP11a; GP11b; neurological disease; genotype; phenotype; diagnosis;
 KW allantoigen; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntingdon's disease; depression; Parkinson's disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..26
 FT /label= signal-peptide
 FT Protein 27..788
 FT /label= mature-protein
 FT Misc-difference 33
 FT /label= I33P
 FT /note= "mutation caused by T192C substitution in
 the coding sequence"
 FN
 XX
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-IB01696.
 XX
 PR 01-OCT-1998; 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K;
 DR WPI; 2000-303801/26.
 DR N-PSDB; AAA09363.

XX Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIIa or GPIIb locus and
 XX determining presence of variant GPIIIa or GPIIb allele or isoform
 PS Claim 16; Fig 4; 55pp; English.

CC Two different forms of GPIIIa, allantoicins PLA1 and PLA2 (for Platelet
 CC Antigen 1 and 2) have been described and can be distinguished using a
 CC monoclonal antibody. The rarer form, PLA2 has sustained a point mutation
 CC at base 192 that causes a nucleotide change from a T to a C and thus a
 CC leucine to proline amino acid substitution at residue position 33.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-Infarct dementia.

SO Sequence 788 AA:

Query Match 99.0%; Score 4112; DB 21; Length 788;
 Best Local Similarity 99.2%; Pred. No. 2e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSGQCLAVSPMCAMCDEALPLGSPRODLKENLKDNCAPESIEFPYSE 60
 DB 27 gpnicttrgvsqcgclavspmcawcsdealpplsprcdlkenllkdncapestiefpyse 86
 QY 61 ARVLEDRPLSDKSGSDSQVTVSPQRLALRLRPDSKNFSIQVRQEDYPVDIYLLMDL 120
 DB 87 arvledrplsdksgsdsgvtvqsprralrlrpdsknfsiqvrqedyvdiyllmdl 146
 QY 121 SYSKMDLMSIQNLGKTLATQMRKLTLSNRIGFARVDRKPVSPYWTISPPELLENPCYDM 180
 DB 147 sysmkdlmsiqnlgtklatqmrkltlsnrlrfgatvdrkpvspypwtisppealennpcydm 206
 QY 181 KTKVCLNMFQKHNLTLDVQTRNEEVRKQSSRNRRDAEGGFDAIMQVDEKIGMKN 240
 DB 207 ktkclnfmfghnltldvqtrneevrkqssrnrndaegegfdaimgatvdekigymkn 266
 QY 241 DASHLDFPTTDAKTHIALDGLAGIVQPNQGCHWGSNDHNSASTTMDYPSGLMTEKLS 300
 DB 267 dashldfpttdakthtaldglagivqpnqgchvgsndhnsasttmdypslgmteklis 326
 QY 301 QKINILFAVTENVVNLQNYSELIRGTTGVLSMDSSNVQLIYDAYGKIRSKVELEVR 360
 DB 327 qkinilfavlenvvnlqnyseilrgttgvylsmdssnvqliydaygkirkveleivr 386
 QY 361 DLPEELSLSPNATCLNNEVTPGLKSCMGLKIDPTVSFSEAVVRGCPQKREKSFRTKPVG 420
 DB 387 dlpeelslspnatclnnevtpglkscmglkidptvsfseavvrpgpqrkksftrtkpvg 446
 QY 421 FKDSLIVQVTFDDCACQAQAEFNSHRCNNGNTEFCGVCRCGPGWLSQCECSEEDYRP 480
 DB 447 fksdlivqvtfddcacqaqaefnshrcnngntefcgcrgpgwlsqcecseddyrp 506
 QY 481 SQODECSPREGQPVCSQGECLCGCCVCHSSDFGKITGTYCECDDFSCVRYRIGEMCSHG 540
 DB 507 sqodecspregqpvcsqgeclcgccvchssdfgkitgttycecddfscvryrigemcsghg 566
 QY 541 QCSGCGICLDSMDTGYVCNCTRTDTCMSSNGLCSGRKCEGSGVCVCIOPSGYDPTCK 600
 DB 567 qcsqcgicldsdmtgyvcnctrttdtcmsnglcsgrkcegsqvcvciopsgydpck 626
 QY 601 CPTCPACTFKEKCEVCKKFDRCALHDENTCNRYCDEIESVVELKDTGKDVNCTYKNE 660
 DB 627 cptcpactfkeckeckkfdreymtntcnrcyrcdesvvelkdtgkdvncnykne 686

QY 661 DDCVVRQYVEDSSGSKILYVEEPCPKGPDLIVLISWGAIIILGALILIKLIT 720
 DB 687 ddcvvrqyvedssgskilyveeepckgpdliyliswgailllgalliklilt 746
 QY 721 IHDRKEFAKFEERAKRAMPDANNPLYKEATSTFTNTYRGT 762
 DB 747 ihdrkefakfeeraraktamplykeatstftntlyrgt 788

RESULT 11
 AAY92452
 ID AAY92452 standard; Protein; 762 AA.

XX AAY92452;
 AC 10-AUG-2000 (first entry)
 DT 10-AUG-2000 (first entry)
 XX GPIIIa variant Arg214Trp.
 KW GPIIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KW allantoicins; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.
 XX MO200020634-A1.
 PN 13-APR-2000.
 PD 01-OCT-1999; 99WC-IB01696.
 PF 01-OCT-1998; 98US-0102624.
 PR (NOVA-) NOVA MOLECULAR INC.
 PA Schappert K;
 PI WPI: 2000-303801/26.

XX Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIIa or GPIIb locus and
 PT determining presence of variant GPIIIa or GPIIb allele or isoform
 PS Disclosure; Page -: 55pp; English.

XX Known polymorphisms in GPIIIa that may be determined to be variants
 CC using methods of the invention include AAY92447-56, which are variant
 CC GPIIIa sequences derived from the mature protein.

CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-Infarct dementia.

CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GPIIIa sequence which appears in Figure 3.

SO Sequence 762 AA:

Query Match 99.0%; Score 4111; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 2.3e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSGQCLAVSPMCAMCDEALPLGSPRODLKENLKDNCAPESIEFPYSE 60
 DB 1 gpnicttrgvsqcgclavspmcawcsdealpplsprcdlkenllkdncapestiefpyse 60
 QY 61 ARVLEDRPLSDKSGSDSQVTVSPQRLALRLRPDSKNFSIQVRQEDYPVDIYLLMDL 120

```

Db 61 arVledrplsksgsdsgyvtqvsppqialrlrpdsskntsiqrvvedypvdiylmld 120
Qy 121 SYSMKDIWSIONTGKTLATQMRKLTSLNRIGFAGVADKRVSYMISPEALENCYDM 180
Db 121 SYSMKDIWSIONTGKTLATQMRKLTSLNRIGFAGVADKRVSYMISPEALENCYDM 180
Qy 181 KTTCLPMFGYKHYLTLDQVTRFENEYKKSQSVNRDAPEGGDAIMQATVCDEKIGWRN 240
Db 181 KTTCLPMFGYKHYLTLDQVTRFENEYKKSQSVNRDAPEGGDAIMQATVCDEKIGWRN 240
Qy 241 DASHLIVFTTDAKTHIALDGLAGIVQPNDOQCHVSDNHYASATTMDYPSLGLMTEKLS 300
Db 241 DASHLIVFTTDAKTHIALDGLAGIVQPNDOQCHVSDNHYASATTMDYPSLGLMTEKLS 300
Qy 301 OKNINLIFAVTENVNLYQNTSELIPTGVLSMSSNVLOLIVAYGKIRSKVELEVR 360
Db 301 OKNINLIFAVTENVNLYQNTSELIPTGVLSMSSNVLOLIVAYGKIRSKVELEVR 360
Qy 361 DLPEELISFNATCLNNEVLPGLKSCMGKIGDTPVSFSIEAKYRGCPQEKESFTTKPVG 420
Db 361 DLPEELISFNATCLNNEVLPGLKSCMGKIGDTPVSFSIEAKYRGCPQEKESFTTKPVG 420
Qy 421 FKSLIVQVTFDDCACAOAEPNSHRNNGNCTFPCGVCRCGPGWLGSQCCESEEDYRP 480
Db 421 FKSLIVQVTFDDCACAOAEPNSHRNNGNCTFPCGVCRCGPGWLGSQCCESEEDYRP 480
Qy 481 SQDECSPREGQPVCSQRCGCLGQCVCVCHSDPFKITGKYCECDDPSCYRYGEMCSGNG 540
Db 481 SQDECSPREGQPVCSQRCGCLGQCVCVCHSDPFKITGKYCECDDPSCYRYGEMCSGNG 540
Qy 541 QCSGDLCDSDMTGYCNCCTRTDTCMSSNGLCSGRKCBGSCVCYIOPSGYGTCEK 600
Db 541 QCSGDLCDSDMTGYCNCCTRTDTCMSSNGLCSGRKCBGSCVCYIOPSGYGTCEK 600
Qy 601 CPICPACATFKKCECKECPKALHDEMTNRCYCDLIESVKELKDTOKDAVNCTYKNE 660
Db 601 CPICPACATFKKCECKECPKALHDEMTNRCYCDLIESVKELKDTOKDAVNCTYKNE 660
Qy 661 DDCVVFQYEDSSGKSILYVVEEPCPKGPDLIVLIVMGAILLIGLAALLIKRLT 720
Db 661 DDCVVFQYEDSSGKSILYVVEEPCPKGPDLIVLIVMGAILLIGLAALLIKRLT 720
Qy 721 IHDRKEPAKFEERARAKWDANPLYKEATSTFTNITYRGT 762
Db 721 IHDRKEPAKFEERARAKWDANPLYKEATSTFTNITYRGT 762

```

RESULT 12

ID AAY92454 standard; Protein; 762 AA.

AC AAY92454;

DT 10-AUG-2000 (first entry)

DE GPIIIa variant Pro407Ala.

XX GPIIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;

XX allantigen; platelet antigen; PLAI; PLI2; Alzheimer's disease; stroke;

XX neurofibromatosis; Huntington's disease; depression; Parkinson's disease;

XX amyotrophic lateral sclerosis; multiple sclerosis; dementia.

XX Homo sapiens.

XX WO200020634-A1.

XX 13-APR-2000.

XX 01-OCT-1999; 99WO-IB01696.

XX 01-OCT-1998; 98US-0102624.

```

PA (NOVA-) NOVA MOLECULAR INC.
XX Schappert K:
XX WPI: 2000-303801/26.
XX
XX Identifying a subject at risk for a neurological disease comprises
XX determination of genotype or phenotype of GPIIIa or GPIIb locus and
XX determining presence of variant GPIIIa or GPIIb allele or isoform
XX
XX Disclosure: Page -: 55pp; English.
XX
XX Known polymorphisms in GPIIIa that may be determined to be variants
XX using the methods of the invention include AAY92447-56, which are variant
XX GPIIIa sequences derived from the mature protein.
XX Identifying a subject at risk for a neurological disease comprises
XX determining the genotype or phenotype of the GPIIIa or GPIIb locus of a
XX subject, and determining the presence of a variant GPIIIa or GPIIb
XX allele or isoform, where the presence indicates an increased risk of
XX neurological disease. The method is useful for identifying a risk,
XX diagnosing and treating a neurological disease comprising Alzheimer's
XX disease, neurofibromatosis, Huntington's disease, depression,
XX amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
XX disease and multi-infarct dementia.
XX Note: This sequence is not given in the specification it was created
XX from the wild type human GPIIIa sequence which appears in Figure 3.
XX
XX Sequence 762 AA:

```

```

Query Match 99.0%; Score 4111; DB 21; Length 762;
Best Local Similarity 99.2%; Pred. No. 2,3e-282;
Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 1 GPNCTTGTGSSCCOCLAVSPMCAMCSDEALPISPRCDLKENLKNCAPESTEFYSE 60
Db 1 gpncttgvssccqclavspmcawcsdealpisprrcdlkenllkncapesstefpse 60
Qy 61 ARVLEDRPLSKSGSDSSQVQVSPORALRLRPDSSKNFSIOVROVEDYVVDIYYLMDL 120
Db 61 arVledrplsksgsdsgyvtqvsppqialrlrpdsskntsiqrvvedypvdiylmld 120
Qy 121 SYSMKDIWSIONTGKTLATQMRKLTSLNRIGFAGVADKRVSYMISPEALENCYDM 180
Db 121 SYSMKDIWSIONTGKTLATQMRKLTSLNRIGFAGVADKRVSYMISPEALENCYDM 180
Qy 181 KTTCLPMFGYKHYLTLDQVTRFENEYKKSQSVNRDAPEGGDAIMQATVCDEKIGWRN 240
Db 181 KTTCLPMFGYKHYLTLDQVTRFENEYKKSQSVNRDAPEGGDAIMQATVCDEKIGWRN 240
Qy 241 DASHLIVFTTDAKTHIALDGLAGIVQPNDOQCHVSDNHYASATTMDYPSLGLMTEKLS 300
Db 241 DASHLIVFTTDAKTHIALDGLAGIVQPNDOQCHVSDNHYASATTMDYPSLGLMTEKLS 300
Qy 301 OKNINLIFAVTENVNLYQNTSELIPTGVLSMSSNVLOLIVAYGKIRSKVELEVR 360
Db 301 OKNINLIFAVTENVNLYQNTSELIPTGVLSMSSNVLOLIVAYGKIRSKVELEVR 360
Qy 361 DLPEELISFNATCLNNEVLPGLKSCMGKIGDTPVSFSIEAKYRGCPQEKESFTTKPVG 420
Db 361 DLPEELISFNATCLNNEVLPGLKSCMGKIGDTPVSFSIEAKYRGCPQEKESFTTKPVG 420
Qy 421 FKSLIVQVTFDDCACAOAEPNSHRNNGNCTFPCGVCRCGPGWLGSQCCESEEDYRP 480
Db 421 FKSLIVQVTFDDCACAOAEPNSHRNNGNCTFPCGVCRCGPGWLGSQCCESEEDYRP 480
Qy 481 SQDECSPREGQPVCSQRCGCLGQCVCVCHSDPFKITGKYCECDDPSCYRYGEMCSGNG 540
Db 481 SQDECSPREGQPVCSQRCGCLGQCVCVCHSDPFKITGKYCECDDPSCYRYGEMCSGNG 540
Qy 541 QCSGDLCDSDMTGYCNCCTRTDTCMSSNGLCSGRKCBGSCVCYIOPSGYGTCEK 600
Db 541 QCSGDLCDSDMTGYCNCCTRTDTCMSSNGLCSGRKCBGSCVCYIOPSGYGTCEK 600

```

QY 601 CPTCPDCTFKKCEVCKEDRGALHDETCNRCRDEIESYKELKTGKDAVNCTYKNE 660
 DB 601 CPTCPDCTFKKCEVCKEDRGALHDETCNRCRDEIESYKELKTGKDAVNCTYKNE 660
 QY 661 DDCVVRFGYEDSSGSKSLIYVVEEPECCKGPDILVLLSVGAILLIGLALLIWKLLIT 720
 DB 661 ddcvvrfgyedsqsksllyvveepcckgpdilvllsvmgaillliglaalllwkllit 720
 QY 721 IHDRKEFAKFEERARAKWDANNPLYKEATSTFTNTITYRGT 762
 DB 721 ihdrkefakfeeararakwdannplykeatsftntityrgt 762

RESULT 13

AAV92455
 ID AAV92455 standard; Protein: 762 AA.

AC AAV92455;

DT 10-AUG-2000 (first entry)

DE GPIIa variant Arg636Cys.

KW GPIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KX allonantigen; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
 KM amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.

PN WO200020634-A1.

PD 13-APR-2000.

PE 01-OCT-1999: 99WO-1B01696.

PR 01-OCT-1998: 98US-0102624.

PA (NOVA-) NOVA MOLECULAR INC.

PI Schappert K;

DR WPI: 2000-303801/26.

PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIa or GPIIb locus and
 PT determining presence of variant GPIIa or GPIIb allele or isoform

PS Disclosure; Page -: 55pp; English.

CC known polymorphisms in GPIIa that may be determined to be variants
 CC using the methods of the invention include AAV92447-56, which are variant
 CC GPIIa sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntingdon's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GPIIa sequence which appears in Figure 3.

XX Sequence 762 AA;

Query Match 99.0%; Score 4111; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 2,3e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTTRGVSSQOCCLAASPMCAWCSDEALPLGSPKCDLKENLNDNCAPESIEEPVSE 60
 DB 1 gpnicttrgvssccqlavspmcawcsdealpLgsprcdlkenlndncapesiefpvse 60
 QY 61 ARVLEDPRLSDKSGSDSSQVTVSPQRIALRLRPDDSKNFSIQVQVEDYPVVIYVLMDL 120
 DB 61 arvledprlsdksqsdssqvtrvspqrialrlrppddsknfsiqvqvedypvviylmdl 120
 QY 121 SYSKKDDLMSIQNLGTYLATQMRKLSNLRIGFCAVYDKPVSEYKISPEALENPECYDM 180
 DB 121 syskkddlmsiqnlgtylatqmrklsnlrigfcafvdkpvsykyispealenpocydm 180
 QY 181 KTTCLPMFGRKHVLTLDQVTRFMEYKVKOSVSNRRAPRGCDATVQATVCEKIGMWRN 240
 DB 181 ktclpmlfgykhvltldqvtrfmevykksvsnrraprgcdatvqatvcekiqwmrn 240
 QY 241 DASHLVEFTTDAKTHIALDGRAGIYVQPNQOCVHSGSDNHSYASTVMDYPSLGLMEKLS 300
 DB 241 dashlvefttdakthialdgragiylvpnqocvhsgsdnhsyastvmidypslglmekls 300
 QY 301 QKNINLIFAVTEENVNLKYONSELIPGTIVGVLSMDSSNVLQILVDAYGKIRSKVELEVR 360
 DB 301 qkninlifavtenvnlkyonselipgtivgvlsmdssnvlqilvdaygkirkvelevr 360
 QY 361 DLPEELISFPAATCLNNEVTPGLKSCMGKRIQDVFSTFEAKVRGCRQEKESTTTPYVG 420
 DB 361 dlpeelistsfpaatclnnevtpglkscmgkriqdvfstfeakvrqcrqekestttpyvg 420
 QY 421 FKDSLIVQVTFPDCCACAOAEPNSHRCNNGNGTFEEGVCVCGPGMGLSCGCECSEEDYRP 480
 DB 421 fkdsliqvtfpdcacacaeapnshrcnngngtfecgvccpgmglscgcecseedyrp 480
 QY 481 SQODECSFREGQPVCSQREGLCGQVCVCHSSDFGKITGKYCECDPFCVRYKGEKSGHG 540
 DB 481 sqodecsfregqpvcsqreqlcgqvcvchssdfgkitgkycecdpfcvrykgeksghg 540
 QY 541 QCSGDCCLCDSDDMTGYCNCITRTDTCMSSNGLLCSGRKCECCSCYCIQPGSGDICEK 600
 DB 541 qcsgdcclcdsddmtgycncitrtdctcmssngllcsgrkceccscyciqpgsgdicek 600
 QY 601 CPTCPDCTFKKCEVCKEDRGALHDETCNRCRDEIESYKELKTGKDAVNCTYKNE 660
 DB 601 CPTCPDCTFKKCEVCKEDRGALHDETCNRCRDEIESYKELKTGKDAVNCTYKNE 660
 QY 661 DDCVVRFGYEDSSGSKSLIYVVEEPECCKGPDILVLLSVGAILLIGLALLIWKLLIT 720
 DB 661 ddcvvrfgyedsqsksllyvveepcckgpdilvllsvmgaillliglaalllwkllit 720
 QY 721 IHDRKEFAKFEERARAKWDANNPLYKEATSTFTNTITYRGT 762
 DB 721 ihdrkefakfeeararakwdannplykeatsftntityrgt 762

RESULT 14

AAV92449
 ID AAV92449 standard; Protein: 762 AA.

AC AAV92449;

DT 10-AUG-2000 (first entry)

DE GPIIa variant Asp118Tyr.

KW GPIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KX allonantigen; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
 KM amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.

PN WO200020634-A1.

PD 13-APR-2000.

XX 01-OCT-1999: 99WO-IB01696.
 XX
 XX 01-OCT-1998: 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K:
 XX WPI: 2000-303801/26.
 XX
 PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIb or GPIIb locus and
 PT determining presence of variant GPIIb or GPIIb allele or isoform
 PS
 PS Disclosure; Page -: 55pp; English.
 CC
 CC Known polymorphisms in GPIIb that may be determined to be variants
 CC using the methods of the invention include AAY92447-56, which are variant
 CC GPIIb sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIb or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIb or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GPIIb sequence which appears in Figure 3.
 CC
 XX
 XX Sequence 762 AA:

Query Match 98.9% Score 4110: DB 21: Length 762:
 Best Local Similarity 99.2% Pred. No. 2.7e-282:
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSQOCLAVSPMCAMCSDEALPLGSPRCDLKNLKNAPESIEFPVSE 60
 DB 1 gpicttrgvsqcgqlavspmcawcsdealpigsprcdlknllkdnapsiefpse 60
 QY 61 ARYLERPLSDKSGDSSQYTOVSFORIALRLRPDDSKNFSIQROVEDYPVDIYLMDL 120
 DB 61 aryledrplsdksqdsqyvtvsforialrlrpdsknfsiqrvgedypvdiylmlyl 120
 QY 121 SYSMKDLMSIONLGTKLTKTORRKLTSLNRIGFAPVOKPSPYVWISPEALENPCYDM 180
 DB 121 sysmkdlmsionlgtkltktorrkltslnrigfapvokpsspyvwispealenpcydm 180
 QY 181 KTTCLPMFGYKHVLTLDVTRFNEEVKKQSYSRNRDAPEGGFDAIMQATVDEKIRGN 240
 DB 181 ktclclpmfgykhvltldvtrfneevkkqsyssrnrdapeggfdaimgatvdekigrn 240
 QY 241 DASHLLVFTTDAKTHIALDGLAGIVQPNQCHVSGSDNHSASTTMDYPSLGLMTEKLS 300
 DB 241 dashllvfttdakthialdglagivqpnqchvsgsdnhsasttmdypsigtmtkls 300
 QY 301 GKNINLIPAVTENYVNLVONYSELIPGTTVGVLSMDSNVQLIYDAIGKIRSKYLEVR 360
 DB 301 gkninlipavtenyvnlyqnyseilpgttvgvlsmdssnvqliydaygkirkylevr 360
 QY 361 DIPPEELISFNATCLNNEVTPGLKSCMGKIDTVSFSIEAKVRCCPEKEKSFIRKPVG 420
 DB 361 dippeelistsfnatclnnevtpglkscmgkldtvsfsieakvrcpkeksfirkpvg 420
 QY 421 FSDSLIVQYTFEDCCACQAQAPNSHRNNGNGTRECVCRCRGPGLWSQCESEEDYRP 480
 DB 421 fksdlivqytfedccacqaqapnshrcnngngtrecvcrcrgpgplwsqecsedyrp 480
 QY 481 SQQDSCSPREGQPVCSQGECLCGCCVCHSSDFGKITKRYCECDDFSCVRRKGEMCSG 540
 DB 481 sqqdcspregqpvcsqgeclcgccvchssdfgkitkrycecddfscvrrkgemcsgh 540

DB 481 sqqdcspregqpvcsqgeclcgccvchssdfgkitkrycecddfscvrrkgemcsgh 540
 QY 541 QCSGDCICLDSWITGYTCNCTTRDTGCMSSNGLCSGKCEGSCVYIQGSIQDCEK 600
 DB 541 qcsgcdicldswitgytcncttrdtgcmssnglcsqgkcegcscvciqgsgydcck 600
 QY 601 CPTCPDCTFFKCEVCEKFKFRGALHDETNCRFCRDEIESEVKEKLDGKDAVACTYKNE 660
 DB 601 cptcpdctfkeceveckfkfrgalhdeyntcnrfcdeiesevkelkdgkdvactykne 660
 QY 661 DDCVVRFOYEDSSGKSLIYVEEPCPKGPDILVLLSVNGAILLGLAALLIWKLLIT 720
 DB 661 ddcvvrfoyedssgksliyyveeepckgpdilvllsvngailllglalllwkllit 720
 QY 721 IHDRKEFAKFEERARAKWDTANNPLVKEATSTFTNTTYRGT 762
 DB 721 ihdrkefakfeerarakwdtannplvkeatstftnttyrgt 762

RESULT 15

ID AAY92453 standard; Protein: 762 AA.

AC AAY92453;

DT 10-APR-2000 (first entry)

DE GPIIb variant Cys374Tyr.

KW GPIIb; GPIIb; neurological disease; genotype; phenotype; diagnosis;

KW allolantigen; platelet antigen; P1A1; P1A2; Alzheimer's disease; stroke;

KW neurofibromatosis; Huntingdon's disease; depression; Parkinson's disease;

KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.

PN WO200020634-A1.

PD 13-APR-2000.

PF 01-OCT-1999: 99WO-IB01696.

PR 01-OCT-1998: 98US-0102624.

PA (NOVA-) NOVA MOLECULAR INC.

PI Schappert K:

DR WPI: 2000-303801/26.

PT Identifying a subject at risk for a neurological disease comprises

PT determination of genotype or phenotype of GPIIb or GPIIb locus and

PT determining presence of variant GPIIb or GPIIb allele or isoform

PS Disclosure; Page -: 55pp; English.

XX Known polymorphisms in GPIIb that may be determined to be variants

XX using the methods of the invention include AAY92447-56, which are variant

XX GPIIb sequences derived from the mature protein.

XX Identifying a subject at risk for a neurological disease comprises

XX determining the genotype or phenotype of the GPIIb or GPIIb locus of a

XX subject, and determining the presence of a variant GPIIb or GPIIb

XX allele or isoform, where the presence indicates an increased risk of

XX neurological disease. The method is useful for identifying a risk,

XX diagnosing and treating a neurological disease comprising Alzheimer's

XX disease, neurofibromatosis, Huntington's disease, depression,

XX amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's

XX disease and multi-infarct dementia.

XX Note: This sequence is not given in the specification it was created

XX from the wild type human GPIIb sequence which appears in Figure 3.

XX Sequence 762 AA:

Query Match	98.98;	Score 4108;	DB 21;	Length 762;
Best Local Similarity	99.28;	Pred. No. 3.7e-282;		
Matches 756; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0

OY	1	GNPNCTT	RGVSSCQOC	LAVSPMCA	MSDEAL	PLGSP	PCDL	KENML	KDNCA	PEST	EPVSE	60																																									
Dp	1	gnpictct	rgvssccqcc	lavspmcawc	sdeaelbrj	gsprcd	kentll	kdnca	pest	ie	fpvse	60																																									
OY	61	ARVLEDR	PLSDKSGSSD	SVOTVO	SPORIAL	RLRP	DDSKNE	SIO	ROVED	PVDI	YYLMDL	120																																									
Dp	61	arvledtr	plsdkgsgssq	vtqvspqr	ialrlr	lpddsk	nefsiq	rvgedy	pvdijyl	mdl		120																																									
OY	121	SYSKKDL	MSIQNL	IGTLAT	QMRK	KLNS	NLRIG	GA	YVDK	VSFY	WTIS	SPPEALE	NTCYM	180																																							
Dp	121	symskddl	msiqnlq	igtlak	qmckl	tanll	igfga	fydk	vpvsy	myl	sppeale	ntpcydm		180																																							
OY	181	KTTCL	PMPGYKH	VLTLD	DOYTR	FRFEE	VEK	KOSYS	SNR	PAR	REG	GDAL	IMQAT	VOCEK	I	GMN	240																																				
Dp	181	kttclpm	pgykhvlt	ldcgvtr	frftee	vkks	vsntda	pegd	ida	mqat	vocek	igmn					240																																				
OY	241	DASHL	LFWT	DAKTH	IALDL	R	L	AGI	VO	PN	GOC	HV	GS	D	NHYS	A	S	AT	MDY	PS	L	G	M	T	E	K	L	S	300																								
Dp	241	dashl	lftwtcak	hialdl	grlagi	vnpdng	qcbv	gsdmh	ysast	ctmd	yps	ig	lme	eks																300																							
OY	301	OKNIN	L	PA	T	EN	V	N	L	X	O	N	S	E	L	I	P	E	T	V	G	L	S	M	O	S	S	N	L	O	L	I	D	A	Y	G	K	I	N	S	K	V	E	L	E	V	R	360					
Dp	301	oknlnl	pa	t	en	v	n	l	x	o	n	s	e	l	i	p	e	t	v	g	l	s	m	o	s	s	n	l	o	l	i	d	y	g	k	i	n	s	k	v	e	l	e	v	r	360							
OY	361	DLPEEL	S	L	S	F	A	N	A	T	C	L	N	N	E	V	I	P	G	L	K	S	C	M	G	L	I	G	D	V	S	R	S	E	A	K	R	G	C	P	O	E	K	E	R	S	F	T	L	K	P	G	420
Dp	361	dipeel	s	l	s	f	a	n	a	t	c	l	n	n	e	v	i	p	g	l	k	s	c	m	g	l	i	g	d	v	s	r	s	e	a	k	r	g	c	p	o	e	k	e	r	s	f	t	l	k	p	g	420
OY	421	FKDSL	I	V	O	V	T	F	D	C	C	A	O	A	E	P	S	H	R	C	N	G	N	G	N	G	N	T	F	E	E	C	G	V	C	R	G	P	G	M	L	S	Q	E	C	S	E	E	D	Y	R	480	
Dp	421	fkdsli	v	o	v	t	f	d	c	c	a	o	a	e	p	s	h	r	c	n	g	n	g	n	g	n	t	f	e	e	c	g	v	c	r	g	p	g	m	l	s	q	e	c	s	e	e	d	y	r	480		
OY	481	SQODEC	S	P	R	E	Q	P	V	C	S	Q	R	G	E	L	C	Q	C	V	C	H	S	S	D	F	E	K	I	T	G	K	I	C	E	D	D	E	S	C	V	A	R	K	E	M	S	G	H	G	540		
Dp	481	sqodecs	p	r	e	q	p	v	c	s	q	r	g	e	l	c	q	c	v	c	h	s	s	d	f	e	k	i	t	g	k	i	c	e	d	d	e	s	c	v	a	r	k	e	m	s	g	h	g	540			
OY	541	QCSG	GD	C	L	C	D	S	D	M	I	G	Y	C	N	C	T	R	T	D	O	M	S	S	N	G	L	C	S	R	G	C	E	G	S	C	V	C	I	O	P	G	S	V	G	P	R	C	E	K	600		
Dp	541	qcsgcd	c	l	c	d	s	d	m	i	g	y	c	n	c	t	r	t	d	o	m	s	s	n	g	l	c	s	r	g	c	e	g	s	c	v	c	i	o	p	g	s	v	g	p	r	c	e	k	600			
OY	601	CPTCP	D	A	C																																																

Search completed: May 19, 2002, 10:38:56
Job time: 9118 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 08:07:51 ; Search time 36.84 Seconds
(without alignments)
505.220 Million cell updates/sec

Title: US-09-673-302a-1
Sequence: 1 GPNICTRGVSSCQCCLAVS.....NNPLYKATSTFTNITYRGT 762

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/2/1aa/5a_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6a_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6b_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6c_COMB.pep:*
7: /cgn2_6/ptodata/2/1aa/6d_COMB.pep:*
8: /cgn2_6/ptodata/2/1aa/6e_COMB.pep:*
9: /cgn2_6/ptodata/2/1aa/6f_COMB.pep:*
10: /cgn2_6/ptodata/2/1aa/6g_COMB.pep:*
11: /cgn2_6/ptodata/2/1aa/6h_COMB.pep:*
12: /cgn2_6/ptodata/2/1aa/6i_COMB.pep:*
13: /cgn2_6/ptodata/2/1aa/6j_COMB.pep:*
14: /cgn2_6/ptodata/2/1aa/6k_COMB.pep:*
15: /cgn2_6/ptodata/2/1aa/6l_COMB.pep:*
16: /cgn2_6/ptodata/2/1aa/6m_COMB.pep:*
17: /cgn2_6/ptodata/2/1aa/6n_COMB.pep:*
18: /cgn2_6/ptodata/2/1aa/6o_COMB.pep:*
19: /cgn2_6/ptodata/2/1aa/6p_COMB.pep:*
20: /cgn2_6/ptodata/2/1aa/6q_COMB.pep:*
21: /cgn2_6/ptodata/2/1aa/6r_COMB.pep:*
22: /cgn2_6/ptodata/2/1aa/6s_COMB.pep:*
23: /cgn2_6/ptodata/2/1aa/6t_COMB.pep:*
24: /cgn2_6/ptodata/2/1aa/6u_COMB.pep:*
25: /cgn2_6/ptodata/2/1aa/6v_COMB.pep:*
26: /cgn2_6/ptodata/2/1aa/6w_COMB.pep:*
27: /cgn2_6/ptodata/2/1aa/6x_COMB.pep:*
28: /cgn2_6/ptodata/2/1aa/6y_COMB.pep:*
29: /cgn2_6/ptodata/2/1aa/6z_COMB.pep:*
30: /cgn2_6/ptodata/2/1aa/6aa_COMB.pep:*
31: /cgn2_6/ptodata/2/1aa/6ab_COMB.pep:*
32: /cgn2_6/ptodata/2/1aa/6ac_COMB.pep:*
33: /cgn2_6/ptodata/2/1aa/6ad_COMB.pep:*
34: /cgn2_6/ptodata/2/1aa/6ae_COMB.pep:*
35: /cgn2_6/ptodata/2/1aa/6af_COMB.pep:*
36: /cgn2_6/ptodata/2/1aa/6ag_COMB.pep:*
37: /cgn2_6/ptodata/2/1aa/6ah_COMB.pep:*
38: /cgn2_6/ptodata/2/1aa/6ai_COMB.pep:*
39: /cgn2_6/ptodata/2/1aa/6aj_COMB.pep:*
40: /cgn2_6/ptodata/2/1aa/6ak_COMB.pep:*
41: /cgn2_6/ptodata/2/1aa/6al_COMB.pep:*
42: /cgn2_6/ptodata/2/1aa/6am_COMB.pep:*
43: /cgn2_6/ptodata/2/1aa/6an_COMB.pep:*
44: /cgn2_6/ptodata/2/1aa/6ao_COMB.pep:*
45: /cgn2_6/ptodata/2/1aa/6ap_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4119	99.2	788	2	US-07-728-215-32
2	4119	99.2	788	4	US-08-938-085A-32
3	3768	90.7	718	1	US-08-444-792-4
4	3768	90.7	718	1	US-08-444-792-4
5	2339	56.3	799	1	US-08-054-077C-2
6	2156	51.9	788	4	US-07-728-215-27
7	2156	51.9	788	4	US-08-938-085A-27
8	1831	44.1	798	2	US-07-728-215-30
9	1831	44.1	798	2	US-08-938-085A-30
10	1676	40.3	577	2	US-07-728-215-29
11	1676	40.3	577	4	US-08-938-085A-29
12	1535.5	37.0	846	4	US-07-728-215-33
13	1535.5	37.0	846	4	US-08-938-085A-33
14	1535.5	36.4	769	2	US-08-789-078-1
15	1535.5	36.4	769	2	US-08-789-078-1
16	1535.5	36.4	769	2	US-08-789-078-1
17	1535.5	36.4	769	2	US-08-789-078-1
18	1535.5	36.4	769	2	US-08-789-078-1
19	1535.5	36.4	769	2	US-08-789-078-1
20	1535.5	36.4	769	2	US-08-789-078-1
21	1535.5	36.4	769	2	US-08-789-078-1
22	1535.5	36.4	769	2	US-08-789-078-1
23	1535.5	36.4	769	2	US-08-789-078-1
24	1535.5	36.4	769	2	US-08-789-078-1
25	1535.5	36.4	769	2	US-08-789-078-1
26	1535.5	36.4	769	2	US-08-789-078-1
27	1535.5	36.4	769	2	US-08-789-078-1

28	1278	30.8	266	5	PCT-US92-10178-3	Sequence 3, Appl1
29	1271	30.6	266	2	US-07-808-457-5	Sequence 5, Appl1
30	1271	30.6	266	5	PCT-US92-10178-5	Sequence 5, Appl1
31	1065	25.6	200	1	US-08-170-596-15	Sequence 15, Appl1
32	883.5	21.3	256	5	PCT-US96-01314-55	Sequence 55, Appl1
33	797	19.2	255	5	PCT-US96-01314-54	Sequence 54, Appl1
34	731	17.8	344	4	US-09-393-554-1	Sequence 1, Appl1
35	731	17.6	164	2	US-07-808-457-19	Sequence 19, Appl1
36	731	17.6	164	4	PCT-US92-10178-19	Sequence 19, Appl1
37	730	17.6	344	4	US-09-393-554-4	Sequence 4, Appl1
38	728	17.5	344	4	US-09-393-554-3	Sequence 3, Appl1
39	726	17.5	344	4	US-09-393-554-5	Sequence 5, Appl1
40	724	17.4	164	2	US-07-808-457-29	Sequence 29, Appl1
41	724	17.4	164	4	PCT-US92-10178-29	Sequence 29, Appl1
42	722	17.4	344	4	US-09-393-554-2	Sequence 2, Appl1
43	688	16.6	251	5	PCT-US96-01314-59	Sequence 59, Appl1
44	656	15.8	150	5	US-07-808-457-17	Sequence 17, Appl1
45	656	15.8	150	5	PCT-US92-10178-17	Sequence 17, Appl1

ALIGNMENTS

RESULT 1
US-07-728-215-32
; Sequence 32, Application US/07728215
; Patent No. 5962643
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Dean
; APPLICANT: Quaranta, Ylto
; APPLICANT: Pytel, Robert
; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
; NUMBER OF INVENTIONS: 43
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,215
; FILING DATE: 19910711
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P31 8717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-728-215-32

Query Match 99.2% Score 4119; DB 2; Length 788;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCQCCLAVSPKAWCSDALPLGSPRCIDKENVLNDNCAPESIEPVS 60
DB 27 GPNICTRGVSSCQCCLAVSPKAWCSDALPLGSPRCIDKENVLNDNCAPESIEPVS 86

```

QY 61 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYVPVDIYYIMDL 120
D 87 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYVPVDIYYIMDL 146
QY 121 SYKMKDDMSIQNLGTLATQMRKLTSLNLRIGFAGVDPKPYSPYMTISPEALENPCYDM 180
D 147 SYKMKDDMSIQNLGTLATQMRKLTSLNLRIGFAGVDPKPYSPYMTISPEALENPCYDM 206
QY 181 KTTCLPMEGYKHVLTLDQVTRFNEEYKOSVSRNRDAPRGFPALIMQATVCEKIGWRN 240
D 207 KTTCLPMEGYKHVLTLDQVTRFNEEYKOSVSRNRDAPRGFPALIMQATVCEKIGWRN 266
QY 241 DASHLVEFTTDAKTHIALDGRLAGIYOPNDGQCHVSDNHYSASTMDYPSLGLMTEKLS 300
D 267 DASHLVEFTTDAKTHIALDGRLAGIYOPNDGQCHVSDNHYSASTMDYPSLGLMTEKLS 326
QY 301 QKINMLIFAVTENVNLYQNTSELIPTGTGVLSMDSSNVLLQILDVAYGKIRSKVELEVR 360
D 327 QKINMLIFAVTENVNLYQNTSELIPTGTGVLSMDSSNVLLQILDVAYGKIRSKVELEVR 386
QY 361 DLPEELISFNATCLNNEVILPGLKSCMGKIGDTVSPFISAKYRGCPQEKESFTTKPVG 420
D 387 DLPEELISFNATCLNNEVILPGLKSCMGKIGDTVSPFISAKYRGCPQEKESFTTKPVG 446
QY 421 FKDSLIVQVTFEDCDCAQOAEPNSHRCNNNGTFEECGVCRGPGWLGSCQECSEEDYRP 480
D 447 FKDSLIVQVTFEDCDCAQOAEPNSHRCNNNGTFEECGVCRGPGWLGSCQECSEEDYRP 506
QY 481 SQODECSPREGQPVCSORGECLGQGCYCHSDGKTKGKCECDTDSVYKKEGMSGSHG 540
D 507 SQODECSPREGQPVCSORGECLGQGCYCHSDGKTKGKCECDTDSVYKKEGMSGSHG 566
QY 541 QCSGDCDCLSDMDWGYCNCCTTRDTQMSNGLLCSGRKCEGSCVCVCIQPGSYGDTCEK 600
D 567 QCSGDCDCLSDMDWGYCNCCTTRDTQMSNGLLCSGRKCEGSCVCVCIQPGSYGDTCEK 626
QY 601 CPTCPDACTFKKEVECKEKKFDRGALHDETCNRYCRDELSYVELDGTGDAVNCTYKNE 660
D 627 CPTCPDACTFKKEVECKEKKFDRGALHDETCNRYCRDELSYVELDGTGDAVNCTYKNE 686
QY 661 DDCVVRQYEDSSGKSLIYVEEPECKPDPDILVLLSWGAILLIGLAALLIMKLLIT 720
D 687 DDCVVRQYEDSSGKSLIYVEEPECKPDPDILVLLSWGAILLIGLAALLIMKLLIT 746
QY 721 IHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNITTYRGT 762
D 747 IHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNITTYRGT 788

```

```

RESULT 2
US-08-938-085A-32
; Sequence 32, Application US/08938085A
; Patent No. 6339148
; GENERAL INFORMATION:
; APPLICANT: Shepard, Dean
; APPLICANT: Quaranta, Vito
; APPLICANT: Pytela, Robert
; TITLE OF INVENTION: A No. 6339148e1 Integrin Beta Subunit and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938, 085A
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/728, 215
; FILING DATE: 11-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parent, Annette S
; REGISTRATION NUMBER: 42,058
; REFERENCE/DOCKET NUMBER: 023070-080210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-938-085A-32

```

```

Query Match 99.2%; Score 4119; DB 4; Length 788;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 GNRICTRQVSSCQOCLAVSPKAMCSDBALPLSGPDCDLKENLKDNCAPESIEFPVSE 60
D 27 GNRICTRQVSSCQOCLAVSPKAMCSDBALPLSGPDCDLKENLKDNCAPESIEFPVSE 86
QY 61 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYVPVDIYYIMDL 120
D 87 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYVPVDIYYIMDL 146
QY 121 SYKMKDDMSIQNLGTLATQMRKLTSLNLRIGFAGVDPKPYSPYMTISPEALENPCYDM 180
D 147 SYKMKDDMSIQNLGTLATQMRKLTSLNLRIGFAGVDPKPYSPYMTISPEALENPCYDM 206
QY 181 KTTCLPMEGYKHVLTLDQVTRFNEEYKOSVSRNRDAPRGFPALIMQATVCEKIGWRN 240
D 207 KTTCLPMEGYKHVLTLDQVTRFNEEYKOSVSRNRDAPRGFPALIMQATVCEKIGWRN 266
QY 241 DASHLVEFTTDAKTHIALDGRLAGIYOPNDGQCHVSDNHYSASTMDYPSLGLMTEKLS 300
D 267 DASHLVEFTTDAKTHIALDGRLAGIYOPNDGQCHVSDNHYSASTMDYPSLGLMTEKLS 326
QY 301 QKINMLIFAVTENVNLYQNTSELIPTGTGVLSMDSSNVLLQILDVAYGKIRSKVELEVR 360
D 327 QKINMLIFAVTENVNLYQNTSELIPTGTGVLSMDSSNVLLQILDVAYGKIRSKVELEVR 386
QY 361 DLPEELISFNATCLNNEVILPGLKSCMGKIGDTVSPFISAKYRGCPQEKESFTTKPVG 420
D 387 DLPEELISFNATCLNNEVILPGLKSCMGKIGDTVSPFISAKYRGCPQEKESFTTKPVG 446
QY 421 FKDSLIVQVTFEDCDCAQOAEPNSHRCNNNGTFEECGVCRGPGWLGSCQECSEEDYRP 480
D 447 FKDSLIVQVTFEDCDCAQOAEPNSHRCNNNGTFEECGVCRGPGWLGSCQECSEEDYRP 506
QY 481 SQODECSPREGQPVCSORGECLGQGCYCHSDGKTKGKCECDTDSVYKKEGMSGSHG 540
D 507 SQODECSPREGQPVCSORGECLGQGCYCHSDGKTKGKCECDTDSVYKKEGMSGSHG 566
QY 541 QCSGDCDCLSDMDWGYCNCCTTRDTQMSNGLLCSGRKCEGSCVCVCIQPGSYGDTCEK 600
D 567 QCSGDCDCLSDMDWGYCNCCTTRDTQMSNGLLCSGRKCEGSCVCVCIQPGSYGDTCEK 626
QY 601 CPTCPDACTFKKEVECKEKKFDRGALHDETCNRYCRDELSYVELDGTGDAVNCTYKNE 660
D 627 CPTCPDACTFKKEVECKEKKFDRGALHDETCNRYCRDELSYVELDGTGDAVNCTYKNE 686
QY 661 DDCVVRQYEDSSGKSLIYVEEPECKPDPDILVLLSWGAILLIGLAALLIMKLLIT 720

```

DB 687 DDCVRFQYEDSSGKSLTYVEEPCPKGPDLLVLLSVMGAILLIGLAALLIMKLIT 746
QY 721 IHDRKEFAKEERARAKNDANNPLYKEATSTFTNITRYGT 762
DB 747 IHDRKEFAKEERARAKNDANNPLYKEATSTFTNITRYGT 788

RESULT 3

US-08-444-792-4
Sequence 4, Application US/08444792
Patent No. 5726037
GENERAL INFORMATION:
APPLICANT: Bodary, Sarah C.
APPLICANT: Gorman, Cornelia M.
APPLICANT: McLean, John W.
APPLICANT: Napier, Mary A.
TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
NUMBER OF SEQUENCES: 18
TITLE OF INVENTION: POLYPEPTIDES
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,792
FILING DATE: 19-May-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380227
FILING DATE: 30-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/218878
FILING DATE: 28-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/821337
FILING DATE: 13-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/444490
FILING DATE: 01-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/290224
FILING DATE: 22-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: P0552PLC304
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-444-792-4

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDPALPSRCDLKENLKNCAPESTIEFPVSE 60
Query Match 90.7%; Score 3768; DB 1; Length 718;
Best Local Similarity 99.3%; Pred. No. 2,5e-295;
Matches 687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DB 27 GPNICTRGVSSCOQCLAVSPMCAMCSDPALPSRCDLKENLKNCAPESTIEFPVSE 86
QY 61 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDSKNFSIOVROVEDYVVDIYVMDL 120
DB 87 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDSKNFSIOVROVEDYVVDIYVMDL 146
QY 121 SYSKMDLWSTONLGTKATQMRKLTSLNRIGFGAFVDPVSPYIYISPPALNPCYDM 180
DB 147 SYSKMDLWSTONLGTKATQMRKLTSLNRIGFGAFVDPVSPYIYISPPALNPCYDM 206
QY 181 KTTCPMFEGYKRVLTLPQVTRFNEEVKKOSVRRRDAPEGGFPAIMQATYCDKIGRN 240
DB 207 KTTCLPMFGYKRVLTLPQVTRFNEEVKKOSVRRRDAPEGGFPAIMQATYCDKIGRN 266
QY 241 DASHLVEFTDAKTHIALDGLAGIVQPNDOCHVSDNHYSASTTMDYPSLGLMTERLS 300
DB 267 DASHLVEFTDAKTHIALDGLAGIVQPNDOCHVSDNHYSASTTMDYPSLGLMTERLS 326
QY 301 QKNMLIFAVTEENVNLQYNSSELIPGTVGLSDSSNVQLYDANGKTRKVELEVR 360
DB 327 QKNMLIFAVTEENVNLQYNSSELIPGTVGLSDSSNVQLYDANGKTRKVELEVR 386
QY 361 DPEELSLSFNATCLNNEYIPGLKSCMGLKIGDYVSFIEAKVRGCPQEKESFTIRPVG 420
DB 387 DPEELSLSFNATCLNNEYIPGLKSCMGLKIGDYVSFIEAKVRGCPQEKESFTIRPVG 446
QY 421 FSDSLIVQYTFEDCCACQAQAEPNSHRCNNGNGTFECGVCRCGPGWLGSCCESEEDYR 480
DB 447 FSDSLIVQYTFEDCCACQAQAEPNSHRCNNGNGTFECGVCRCGPGWLGSCCESEEDYR 506
QY 481 SODDCSPREGQPVCSORGECLCGOCVCHSSDFGKITKCYECDDFVSVRKKGECSCSHG 540
DB 507 SODDCSPREGQPVCSORGECLCGOCVCHSSDFGKITKCYECDDFVSVRKKGECSCSHG 566
QY 541 QCSGCDLCLDSMTGYNCCTRTDTCMSNGLCSGRGKEGSCGVCYC10P6SYGDTCEK 600
DB 567 QCSGCDLCLDSMTGYNCCTRTDTCMSNGLCSGRGKEGSCGVCYC10P6SYGDTCEK 626
QY 601 CPTCPDCTFEKCEYCKEFGALHDENTCNRYCRDEIESYKELKDTGKDAVNCTYNE 660
DB 627 CPTCPDCTFEKCEYCKEFGALHDENTCNRYCRDEIESYKELKDTGKDAVNCTYNE 686
QY 661 DDCVRFQYEDSSGKSLTYVEEPCPKGPD 692
DB 687 DDCVRFQYEDSSGKSLTYVEEPCPKGPD 718

RESULT 4

US-08-445-042-4
Sequence 4, Application US/08445042
Patent No. 5726290
GENERAL INFORMATION:
APPLICANT: Bodary, Sarah C.
APPLICANT: Gorman, Cornelia M.
APPLICANT: McLean, John W.
APPLICANT: Napier, Mary A.
TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,042

```

FILING DATE: 19-May-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380227
FILING DATE: 30-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/218878
FILING DATE: 28-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/821337
FILING DATE: 13-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/444490
FILING DATE: 01-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/290224
FILING DATE: 22-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: P05522PIC3D1
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-445-042-4

```

```

Query Match          90.7%: Score 3768; DB 1; Length 718;
Best Local Similarity 99.3%; Pred. No. 2.5e-295;
Matches 667; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSPRDLKENLKNCAPESEIEPVE 60
DB 27 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSPRDLKENLKNCAPESEIEPVE 86
QY 61 ARVLEDRPLSDKSSGSSQVTVQSPORIALRLAPDDSKNFSIOVROVEDYVDYIYLM 120
DB 87 ARVLEDRPLSDKSSGSSQVTVQSPORIALRLAPDDSKNFSIOVROVEDYVDYIYLM 146
QY 121 SYSKMDLMSIQNLGTRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALENFCY 180
DB 147 SYSKMDLMSIQNLGTRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALENFCY 206
QY 181 KTTCLPMFGYKHVLLTDOVTRFNEEVKROSVSRNDAEFGFDALIMQATVDEKIGWN 240
DB 207 KTTCLPMFGYKHVLLTDOVTRFNEEVKROSVSRNDAEFGFDALIMQATVDEKIGWN 266
QY 241 DASHLLVFTTDAKTHIALDGRLAGIYQPNDDGCHVSDHHYSASTMDYPSLGLMTEKLS 300
DB 267 DASHLLVFTTDAKTHIALDGRLAGIYQPNDDGCHVSDHHYSASTMDYPSLGLMTEKLS 326
QY 301 QKNINILFAVTENVNMLYONYSSELIPGTVGVLSMDSNVLLIYDAYKISKVELEYR 360
DB 327 QKNINILFAVTENVNMLYONYSSELIPGTVGVLSMDSNVLLIYDAYKISKVELEYR 386
QY 361 DLPEELSLSFNATCLNNEVYIPGLKSCMGLKIGDVSFSIEARVRCGPQEKESFTIKPV 420
DB 387 DLPEELSLSFNATCLNNEVYIPGLKSCMGLKIGDVSFSIEARVRCGPQEKESFTIKPV 446
QY 421 FDSLLVQYTPFDCCACQQAEPNSHRCNNGNGTFFECGVCRCRGPGLSGCCSEEDYR 480
DB 447 FDSLLVQYTPFDCCACQQAEPNSHRCNNGNGTFFECGVCRCRGPGLSGCCSEEDYR 506
QY 481 SODDECSPREGQPVCSQRECECLGCGVCVCHSSDFGKITGKCCDDPSCVRYKGEWCSGHG 540
DB 507 SODDECSPREGQPVCSQRECECLGCGVCVCHSSDFGKITGKCCDDPSCVRYKGEWCSGHG 566

```

```

QY 541 QCSGDCICDSDWMTGYCNCITRTDTOMSSNGLLSGKGEKCGSCVCIOGSGYDTCEK 600
DB 567 QCSGDCICDSDWMTGYCNCITRTDTOMSSNGLLSGKGEKCGSCVCIOGSGYDTCEK 626
QY 601 CPTCPDACEFKKECECKKFKDGRGALHDETCNRCYCRDEIESYKELDKDGAVNCTYKNE 660
DB 627 CPTCPDACEFKKECECKKFKDGRGALHDETCNRCYCRDEIESYKELDKDGAVNCTYKNE 686
QY 661 DDCVVRFOYEDSSGSKSLIYVEEPCPKGPD 692
DB 687 DDCVVRFOYEDSSGSKSLIYVEEPCPKGPD 718

```

RESULT 5

```

US-08-054-077C-2
Sequence 2, Application US/08054077C
Patent No. 5527679
GENERAL INFORMATION:
APPLICANT: HEMLER, MARTIN E.
APPLICANT: RAMASWAMY, HEMAVATHI
TITLE OF INVENTION: HUMAN INTEGRIN 5 SUBUNIT PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
ADDRESS: CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/054,077C
FILING DATE: 27-APR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/694314
FILING DATE: 01-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, DAVID S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 40937
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-054-077C-2

```

```

Query Match          56.3%: Score 2339; DB 1; Length 799;
Best Local Similarity 55.5%; Pred. No. 3.9e-180;
Matches 433; Conservative 118; Mismatches 205; Indels 24; Gaps 11;

```

```

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSP-----RODLKENLKNCAPESE 55
DB 24 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSP-----RODLKENLKNCAPESE 80
QY 56 PVESEARVLEDRPLSDKSSGSS--QVTVQSPORIALRLAPDDSKNFSIOVROVEDYVDYI 114
DB 81 SPASSFHLRLPLSSKSSGSSAGWDVYQMTQELAVNLROGDKTTPOLQRYVEDYVDYI 140
QY 115 YLMDLSYSKMDLMSIQNLGTRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALE 174
DB 140 YLMDLSYSKMDLMSIQNLGTRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALE 174

```



```

Db 141 YLMDLSLMSKDDLDNIRSLGFKLAEEMKRLTNSFRLGFGSVFKDIDSPFSY-TAPRYOT 199
QY 175 NPC--YDMKTCCLPMFGYKHYVLTLDQVTRFNEVKKQSVSRNRDAPBEGCFDAIMQATVC 232
Db 200 NPIGKYLFPNCVYSGFGRHLPLTDVDSFNEEVKROKVRNDRAPBEGFPAVLQAAAC 259
QY 233 DEKIGMRNDASHLVFTTDAKTHIALDGRAGIYQPNDOCHGSDNNH SASFTMDYPSL 292
Db 260 KEXIGMRNDASHLVFTTDAKTHIALDGRAGIYQPNDOCHGSDNNH SASFTMDYPSL 319
QY 293 GLMTEKLSQKINILFAVTENVNLYONYSELIPGTVGVLSMDSNNVQLIVDAYKIR 352
Db 320 ALLGKLAEMNNILFAVTENVNLYONYSELIPGTVGVLSMDSNNVQLIVDAYKIR 379
QY 353 SKVELEVRDLPEELSLSFNATCLINNEYIPQLKSCMGLKIGTVSFIKAVRGP-QEKE 411
Db 380 SKVELEVRDLPEELSLSFNATCLINNEYIPQLKSCMGLKIGTVSFIKAVRGP-QEKE 439
QY 412 KSTFIKPVGFKDSLIVQVFDCCACQAOAEPNSHRCNNGNGFEGCVGRGPGWLGSGC 471
Db 440 HVFLARVRGERSLLEVGVYVNTGCGSVGLEPNSARC-NGSGTYVGLCEGSPGYLGTIC 498
QY 472 ECSEEDYRPSQDECSFREGQPVCSQREGCLCGQCVCHSSDFGKITGKYCECDPFCVRY 531
Db 499 ECQDGENQSYQNLKRAEAKPLCSGRGDCSCNOCSEFESEFKITGPFCECDNFSCARN 558
QY 532 KGEKSGHGQSCGDCDCLDSMTGYVCNCTRTDTCMSNGLLSGRGKCEGSCVCIOP 591
Db 559 KGVLCSGHGEGCECHAGYIGDNOCNSTDITCRGRDQICSRHGLCGOCCTEP 618
QY 592 GSYGDTCEKPTCPDCTFKECEVECKKPFGRGLHDENTCNRCOREIBS-VKELKDTGK 650
Db 619 GAGGEMCEKPTCPDCTFKECEVECKKPFGRGLHDENTCNRCOREIBS-VKELKDTGK 677
QY 651 DAVNCTFKNDDCVARQYVEDSSKSIIVVEEPCPKPDILVLLVSMGAILLIGLA 710
Db 678 EAVLCFKTKAKDCVMFTYELPSGSKNLTVLREPEGNTFPMATLILAVGSLILVGLA 737
QY 711 ALLIWKLLITHDRKEFAKFEERARAKMDTANPLYKATSTFT-----NITYRGT 762
Db 738 LLAIWKLVLTHDRKEFAKFEERARAKMDTANPLYKATSTFT-----NITYRGT 797

RESULT 6
: Sequence 27, Application US/07728215
: Patent No. 5962643
: GENERAL INFORMATION:
: APPLICANT: Sheppard, Dean
: APPLICANT: Quaranta, Vito
: APPLICANT: Pytel, Robert
: TITLE OF INVENTION: A NO. 5962643el Integrin Beta Subunit and Uses
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: United States of America
: Zip: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07728, 215
: FILING DATE: 19910711
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815

```

```

: REFERENCE/DOCKET NUMBER: P31 8717
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-8949
: INFORMATION FOR SEQ ID NO: 27:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 788 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-728-215-27

Query Match 51.98: Score 2156; DB 2: Length 788;
Best Local Similarity 50.68; Pred. No. 2,1e-155;
Matches 389; Conservative 133; Mismatches 231; Indels 16; Gaps 8;

QY 1 GNICTTR-----GVSSQCCLAVSPMCAMCSDAL--PLC-SIPKDLKENILKDNCAPE 52
Db 14 GRNDSRTNWLCLGAEFTCEDCLLIGPOCAMCAQENFTPSGVEGRCDTPANLLAKGCOLN 73
QY 53 STEFVSEARVLEDRPLSDKSGSSQVTVQSPORIALRLRPDSKNFSIOVROVEDYPV 112
Db 74 FIENPVQVEILKKNPLSVGRKNSSDIYOIAPOSLLIKLRPGCAQOTLVHVROTEDYPV 133
QY 113 DLYYLMDSLMSKDDLNSTONIGTYLATOMKRLTNSFRLGFGSVFKDIDSPFSY 172
Db 134 DLYYLMDSLMSKDDLNSTONIGTYLATOMKRLTNSFRLGFGSVFKDIDSPFSY 192
QY 173 LENPCYDMKTTCLPMFGYKHYVLTLDQVTRFNEVKKQSVSRNRDAPBEGCFDAIMQATVC 232
Db 193 IANPCSSIFPECLPFFGKRIHPLTNDAREFNEYKRNKISANIDTPGCGDALIMQATVC 252
QY 233 DEKIGMRNDASHLVFTTDAKTHIALDGRAGIYQPNDOCHGSDNNH SASFTMDYPSL 292
Db 253 KEXIGMRNDASHLVFTTDAKTHIALDGRAGIYQPNDOCHGSDNNH SASFTMDYPSL 312
QY 293 GLMTEKLSQKINILFAVTENVNLYONYSELIPGTVGVLSMDSNNVQLIVDAYKIR 352
Db 313 GOLIDKLVQNNVLLFAVYQEQVLYENYAKLIPATVGLLOKDSGNILOLISAYEELR 372
QY 353 SKVELEVRDLPEELSLSFNATCLINNEYIPQLKSCMGLKIGTVSFIKAVRGP-QEKE 412
Db 373 SEVELEVRDLPEELSLSFNATCLINNEYIPQLKSCMGLKIGTVSFIKAVRGP-QEKE 431
QY 413 SFTIKPVGFKDSLIVQVFDCCACQAOAEPNSHRCNNGNGFEGCVGRGPGWLGSGC 472
Db 432 HIIRKPVGLDALLELVSPENCDCQKEVEVNSKCHHGNGSFQCGVACHPHMGPRCE 491
QY 473 CSEEDYRPSQDECSFREGQPVCSQREGCLCGQCVCHSSDFGKITGKYCECDPFCVRY 532
Db 492 CGED--MLSTDSCKEADPHPSGSRGDCYCGQCIHLSIPGNTIGPYCQCDNFSVVRK 548
QY 533 GEMKSGHGQSCGDCDCLDSMTGYVCNCTRTDTCMSNGLLSGRGKCEGSCVCIOPG 592
Db 549 GLLCGNNGDCDEGECVCRGWTGVCNCTRTDTCMSNGLLSGRGKCEGSCVCIOPG 608
QY 593 SYGDTCEKPTCPDCTFKECEVECKKPFGRGLHDENTCNRCOREIBS-VKELKDTGK 651
Db 609 ASGPCTCERPTCPDCTFKECEVECKKPFGRGLHDENTCNRCOREIBS-VKELKDTGK 666
QY 652 AVNCTYKNDDCVARQYVEDSSKSIIVVEEPCPKPDILVLLVSMGAILLIGLA 711
Db 667 SVSCSLQGENECLLFTTLDNEGKTIHISINEKCPKPPNIPMILGSLATLIGVLL 726
QY 712 LLAIWKLVLTHDRKEFAKFEERARAKMDTANPLYKATSTFTFTYR 760
Db 727 LCIWKLVSFHDRKEFAKFEERARAKMDTANPLYKATSTFTFTYR 775

RESULT 7
: US-08-938-085A-27
: Sequence 27, Application US/08938085A

```

```

Patent No. 6339148
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pytela, Robert
TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Theoreof
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,085A
FILING DATE: 26-SEP-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,215
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parent, Annette S.
REGISTRATION NUMBER: 42,058
REFERENCE/DOCKET NUMBER: 023070-080210DS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-938-085A-27

```

```

Query Match 51.9%; Score 2156; DB 4; Length 788;
Best Local Similarity 50.6%; Pred. No. 2.1e-165;
Matches 389; Conservative 133; Mismatches 231; Indels 16; Gaps 8;

QY 1 GPNICITR-----GVSSCOCLAVSPMCMSDEAL--PLG--SPRDLKENLTKNCAP 52
DB 14 GRNDSRTMILCLGGAETCEDCLIGQCAMCAOENFTPHSGVGECDDTPANILAKCQ 73
QY 53 SIEFPVSEARVLEDRPLSDKSGSSQVQVSPORIALRLRPDDSKNFSGIOVROVEDYV 112
DB 74 FIENPVSOYEILKNRKLPSVGRKNSSDIYQIAQSLILKLRGAGOTLOVHRQEDYV 133
QY 113 DIYVLDLSYMKDLSIWNIGTKLATQMRKLTSLNRIGFAGFVDKVPSPMYTSPPEA 172
DB 134 DIYVLDLSASMDDDLTKEIGSLSKEMSKLTSNFRIGFSGFVKKPSPVYK--TTPE 192
QY 173 LBNPCVDMKTTCLPMFGYHVLTLTQVVRNNEVYKQSVSKRRDAPREGFPAIMQATVC 232
DB 193 IANPCSSIFPCLPTEGFKHILPLTNDARFNEIYKNOKISANIDPEGGFAIMQAAC 252
QY 223 DEKIGRNDASHLLVFTTAKTHIALDGRLAGIYQVNDGCHVGSNDHVSASTMDYPSL 292
DB 253 KEKIGRNDASHLLVFTTAKTHIALDGRLAGIYQVNDGCHVGSNDHVSASTMDYPSL 312
QY 293 GLMTERLSQKNINLIFAVTEENVVNYQYSELIPGTVGLSMDSSNVLTQIVDYGKIR 352
DB 313 GOLIDKLVQNNVLLIFAVTQEOVHYLEYAKLIPGATVGLLQKDSGNILQLTISAYEELR 372
QY 353 SKVELEVRDLPELISFNATCLNNEVIPGLKSCMGLKIGTVVPSISIAKVRGCPQOEK 412

```

```

DB 373 SEVELEVLGDTGELINLFTAIQNNGLTLFHOHKKSCSHMKVGDIAFSVYVNIPIHC-ERRSR 431
QY 413 SFTIKPVGPKDSLIVQVFPEDCCOAOAEPHSHSCNNGNGFEGYCRGCPGMIGSCE 472
DB 432 HIIRKPVGLDLELVISPECNDCQKEVENSSCHHGNSFCQGVACHPGHMPRCE 491
QY 473 CSEEDYRPSQODECSPREQPVCSORGECILGQGCYCHSDGFKITGKICEEDDFSCVYRK 532
DB 492 CGED---MISTSCKEAPDPHSCSGRGCYCGQCCILHSPIGNITGPIYCGQDNFNSCVYRK 548
QY 533 GEMSGHGGCSCGDCCLDSBWTGYICNCTRTDTOMSSNGLICSGRGKCEGSCVCIQPG 592
DB 549 GLICGNGDCDCGECVCHSGGTGECVCTTSDVCHSDVGLCSRGDCVCGKCVCTNPG 608
QY 593 SYGDTCEKCPGCPDACCTFKKECKKRDGALHDEMTCNRYCROEISVYELMDTGKD- 651
DB 609 ASGPTCEKCPGCPDACCTFKKECKKRDGALHDEMTCNRYCROEISVYELMDTGKD- 666
QY 652 AVNCTYKNEDDCVRFQYEDSSGSKILYVEEPECPKGPDLVLLSVMAILLIGLAA 711
DB 667 SVCSLQGENECLIFLTITDNECKTIIHSINEXDCPKRPNIPMIMLGVSLATILIGVVL 726
QY 712 LIAKLLITIHDKFEPAFEERARAKMDTANNPLYKATISTFTNITVR 760
DB 727 LCIWKLVSFHDRKEVAKFEARSKAKWGTGNPLYRSTSTFTNVTYK 775

```

```

RESULT 8
US-07-728-215-30
Sequence 30, Application US/07728215
Patent No. 5962643
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pytela, Robert
TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Theoreof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESS: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 798 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-728-215-30

```

```

Query Match 44.1%; Score 1831; DB 2; Length 798;
Best Local Similarity 44.9%; Pred. No. 3.3e-139;
Matches 350; Conservative 134; Mismatches 268; Indels 28; Gaps 14;

```


Db 620 CTPKFEQDQVCEMCOQTLGVCABHKEVCQCRAFNKEKRD--TCIOECYSFNITKVESRD 677
 QY 644 EL-KDTGKDAV-NCTYKNEEDCVVRFOYEDSGSKSLIYVEEPKPGKPDILVLLSW 701
 Db 678 KLPQVPDPVPSHCKEKEDVDCCMFYFISVNGNVEVAVHVENPECTGPDILPIYAGV 737
 QY 702 GAILLIGLALLIMKLLITIHDRKEFAFEERARAKWDTANNPLYKEATSFPTNTYNG 761
 Db 738 AGIVLIGLALLIMKLLIMIHDRREFAKFEKEMAKWDTGENPIKSAVTIVNKRKYS 797

RESULT 10

US-07-728-215-29
 ; Sequence 29, Application US/07728215
 ; Patent No. 5962643
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Dean
 ; APPLICANT: Quaranta, Vito
 ; APPLICANT: Pytela, Robert
 ; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/728,215
 ; FILING DATE: 19910711
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 577 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-728-215-29

Query Match 40.3%; Score 1676; DB 2; Length 577;
 Best Local Similarity 51.1%; Pred. No. 6,5e-127;
 Matches 297; Conservative 102; Mismatches 174; Indels 8; Gaps 5;

QY 121 SYSMODMSIONLGTOKRATOKRKLTSNRIFGFAVDKPVSPYATISPEELNPPCIDM 180
 Db 1 SASMDLDTLTKELSGLSKEMSKLTSNRIIGFSGVEXKVPSPFK--TTPPEIAMPCCSI 59
 QY 181 KTTCLPMFGYKHLVLTIDVTRFNEVEKKOSVSRNDAPEGGFDALMOATVDEKIGRN 240
 Db 60 PYICPTPEFKHILPLTNAERFNEIVKOKISANIDNPEGSDIMQAAVCCKEKIGRN 119
 QY 241 DASHLLVFTTDAKTHIALDGLAGIVQPNQGCCHVSDNHVSASTTMDYPSLGLTKELIS 300
 Db 120 DSHLLVFSVDSHSGMDSKLAGIYIPNDGLCHLDSKNEYSMSIVMEVPTIGQLIDKVV 179
 QY 301 QKNINILPAVTENVNVLQNIQNSSELLPGTIVGLSDSSNVTLQIYDAVGKIRSKVELEVR 360

Db 180 QNNVLLIFAVTQEQVPLYENYAKLIPGATVGLLHKDSNNIQLIISAVEELRSEVELEVL 239
 QY 361 DLPELISLSEFNATCLNNEVILPGLSKMGLKIGDTPSFSLIAKVRCCPEKRSFTIRPVG 420
 Db 240 GTEGLINTSEFSAVCNNGTLFPHQKCLHMKVGETASFVNTVSIHNC-FRKRHYIIRKVG 298
 QY 421 FNDSLIVQVTFPCDCACQAOAEPNSHRCNNGNGFECCVCRGPAWLSQCECESEEDRP 480
 Db 299 LQDTLELLVSPSCSDQCKEYEVNSKCHNGSGVQCVACNNGHMRPHECEGDTL-- 356
 QY 481 SQQDCSPREGQPVCSQGBELCGQCVCVHSSDFEKTITKYECDPFCVRRKMGKSGHG 540
 Db 357 -STDSCKETPDHPSCSGRGDVCYGCQICHLSPYGNIVGPYCCDNFSCVRRKGLLGDNG 415
 QY 541 QCSGDCICDPSMWTGVCNCTTRDTCMSNGLCSGKCEGSCVCVIGQSGYDTCR 600
 Db 416 DECEGCVCRSGMTGTCNCTTSTDTCLSEGTCLSGRGDVCYGCVCVTNAGAGPTCE 475
 QY 601 CPTCPDCTFKKEVCECKKFRGALHDNFCNRCRDEIESVKELKPDGKD-AVNCYKN 659
 Db 476 CPTCSDPCNSKRSCTIECHLSADQPGGE--CYDKCKLAGVTSKEADFSKDSVSCSIQ 533
 QY 660 EDDCVRFQYEDSSGKSLIYVEEPKPGKPDILVLLSV 700
 Db 534 ENECLITFLISTDNEGKTIHINISEKDCPKPIMPIMLGV 574

RESULT 11

US-08-938-085A-29
 ; Sequence 29, Application US/08938085A
 ; Patent No. 6339148
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Dean
 ; APPLICANT: Quaranta, Vito
 ; APPLICANT: Pytela, Robert
 ; TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/938,085A
 ; FILING DATE: 26-SEP-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/728,215
 ; FILING DATE: 11-JUL-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parent, Annette S.
 ; REGISTRATION NUMBER: 42,058
 ; REFERENCE/DOCKET NUMBER: 023070-0802100S
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 577 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-938-085A-29

Query Match 40.3%; Score 1676; DB 4; Length 577;
 Best Local Similarity 51.1%; Pred. No. 6.5e-127;
 Matches 297; Conservative 102; Mismatches 174; Indels 8; Gaps 5;

QY 121 SYSKMDLSIONLGTATKATORKLTSNRIFGFAVYKPVYATISPEALNCPDLM 180
 DB 1 SASMDLWIKELGSLSKSKSNTFRIFGFEVPEVSPFK-TTPEIAPNCSS 59
 QY 181 KTKLPMFGKHYLTLDVTRNEEVEKQSVSRNDAPEGFDIMQATVDEKIGMN 240
 DB 60 PYCLPTEFGKHLPLTNAERFNEVYKOKISANDNEGGFDIMQAAVCEKIGMN 119
 QY 241 DASHLLVETTDKATIALDGLAGIVQPDGQCHVSDNHYASTTMDYPSLGLMTEKIS 300
 DB 120 DSHLLVETTDKATIALDGLAGIVQPDGQCHVSDNHYASTTMDYPSLGLMTEKIS 179
 QY 301 OKNINLIFAVTEVYVLYXONYSLLPGTVGLSMDSNVLLIDAYGKIRSKYLEVR 360
 DB 180 ONNVLLIFAVTOEYVLYXONYSLLPGTVGLSMDSNVLLIDAYGKIRSKYLEVR 239
 QY 361 DLPEELSLFNATCLNNEYIPLGLSKCMGLKIGDVSFSTIAKVRGCPQEKESFTIKPYG 420
 DB 240 GDEGLNTSFAVNCNGTLEPHQKCLHMKVGTASFNTVSIPLNC-EKSRHVITIKPYG 298
 QY 421 FKDSLIVQVTPDCACQAOAEPNSHRCNNGCTFECGYCRGCPGLGSCQCESEEDYR 480
 DB 299 LGDTLILVSPESCSCCQKQEVENSSKCHNGNSYQGCACNPFHMGPHCEGEDTL-- 356
 QY 481 SQODESPRAGOVPCSORGECLOGCVCHSDPGKITGVCYECDDPCSVRYGEMSGHG 540
 DB 357 -STDCKETPDHPSGSGRGCTGCTCHLSPIGNITGYCQCDNCSVRHNGLLGCDNG 415
 QY 541 QCSGDLCDSDMTGYGNCRTPTDCMSNGLCSGRKCEGSCVCIOPSGYDGTCP 600
 DB 416 DCCGEGVCSGWTGXCNCTSTPCISBDGTLCSGRDVCYKCTNPGASGPTCR 475
 QY 601 CPTCPDCTKKECEVCKEFDKRALDENTCNRYCDELESYELKDGTGD-AVNTYKN 659
 DB 476 CPTCSPCNKSRKICHSADQPGEE-CYDKCLAGVITISKEDFSKSSVCSLGG 533
 QY 660 EDDCVAFQYEDDSGKSLYVEEPECPKPDILVLLSV 700
 DB 534 ENCLITFLSTNEKGTITIHNSKDCPKRPPIPMIMGV 574

RESULT 12
 US-07-728-215-33
 : Sequence 33, Application US/07728215
 : Patent No. 5962643
 : GENERAL INFORMATION:
 : APPLICANT: Sheppard, Dean
 : APPLICANT: Quaranta, Vito
 : APPLICANT: Pytela, Robert
 : TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
 : THEREOF
 : NUMBER OF SEQUENCES: 43
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
 : STREET: 4370 La Jolla Village Drive, Suite 700
 : CITY: San Diego
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 92122
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/728, 215
 : FILING DATE: 19910711
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P31 8717
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-8949
 TELEFAX: (619) 535-9001
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 846 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear

US-07-728-215-33

Query Match 37.0%; Score 1535.5; DB 2; Length 846;
 Best Local Similarity 38.7%; Pred. No. 2.3e-115;
 Matches 320; Conservative 128; Mismatches 292; Indels 87; Gaps 21;

QY 4 ICTTRGVSSCOCLAVSPMCASDEALPLGSPHCDLKENLKNCAPESTIEPPVSEARV 63
 DB 37 VSTCASKKCHTCTQTEG-CAMCQDPFK-GQSRCTYNTSSL--CPEFAVSPITVEOI 91
 QY 64 LEDRPL-----SDKSGDSS----- 78
 DB 92 LVNKNLTQYAEIAGGAGGAGSGSSSSSYSSSSSSSPYSSSSSSSASAGTEYSAG 151
 QY 79 QVTQVSPORIALRLPDDSKNFSIOVRVEDYPVDIYVLYMDLSYMKDLSIONLGTKL 138
 DB 152 EIVQIQPSMLALRVNKNHNIKISYQAEYVPDYLYMDLSYMKDLSIONLGTKL 211
 QY 139 ATQKRLTSNLRIGFAVYKPVSPYATISPEALNCPYDMKTCCLPMFGKHYLTLD 198
 DB 212 SETMRITNNHFLGSGSVFDVLMRYV-STIPKLEHPC---DNCKAPYQOHPMLNN 266
 QY 199 QVTRFNEVEKQSVSRNDAPEGFDIMQATVDEKIGMNDAHSLVFTTDAKTIAL 258
 DB 267 NTESFNSNKNATYVSGNIDAPGEGFDIMQATVDEKIGMNDAHSLVFTTDAKTIAL 326
 QY 229 DGLRGIQVQPDGQCHVSDNHYASTTMDYPSLGLMTEKLSQNNINLIFAVTEVNVLY 318
 DB 327 DGLRGIQVQPDGQCHVSDNHYASTTMDYPSLGLMTEKLSQNNINLIFAVTEVNVLY 386
 QY 319 QVTSLEIGCTYVGLSMDSNVLLIDAYGKIRSKYLEVRDLPEELSLFNATCLNE 378
 DB 387 EKLVEHIGSSSAKLDNDSSNVLEKKEIRKISSVEMK-DNATGDKITTFSSCLSN 445
 QY 379 VPLGLSKCMGLKIGDVSFSTIAKVRGCPQEKESFTIKPYGFKDSLIVQVTPDCD 435
 DB 446 PEVQTSKCDNLKEGQVAFVFTAOIOLKCPEDRDMTORIHISPIGNEVMQIOLMLCSC 505
 QY 436 ACQ-----AQAEPSNHRNNGNGTFCGVCRCGFGMLGSCQCESEEDYRPSQDECS 489
 DB 506 PCENSGSIGYQVQANS--C-SGHGTSKMGICNCDDSYGNGNCECATDLSKSFANDYSCR 562
 QY 490 EGQPV---CSQGBELCQCVCHS--SDPGKITGVCYECDDPCSVRYGEMSGHG--HGQ 542
 DB 563 ADSTSTDSGSGHCCVACBCHKRPNIETIISGHCHCDNPSCRNNNQLCSGDHGTG 622
 QY 543 SCGDCLDSDDMTGYGNCCTTPTDCM--SSNGLCSGRKCEGSCVCI--QPSGYDGTG 598
 DB 623 ECGRCCKRGWGTSGNCGQSDNDTCMPGCGGEGICGHGTCEGVCYKCIYNQGRSGRHC 682
 QY 599 EKCPTCPDCTKKECEVCKEFDKRALDENTCNRYCDE---ETESYELKDGTGDAVN 654
 DB 683 EKCPTCSGRQELKDCVOCOMYKTELKNGDCARNCQFVGVGEV-EIDET-KDEOM 740
 QY 655 CTKNEDCVAFQYEDDSGKSLYVEEPECPKPDILVLLSVKMAIILIGLAALI 714
 DB 741 CFEDEDDCKEMFKYSE--QGLHYVAQENKCECPKPVMLGIVGVIAIYVGLAILL 798
 QY 715 WKLLTTHDRKEFAFEERARAKWDTANNPYKATSTFTNITYRG 761

Db 799 WKLLTTHDRREFARFEKERMANAKMDTGENDPIYKQATSTFKNPMYAG 845

RESULT 13

US-08-938-085A-33
Sequence 33, Application US/08938085A
Patent No. 6339148
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pyeala, Robert
TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Theoreof
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,085A
FILING DATE: 26-FEB-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/728,215
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parent, Annette S.
REGISTRATION NUMBER: 42,058
REFERENCE/DOCKET NUMBER: 023070-080210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 846 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-938-085A-33

Query Match 37.0%; Score 1535.5; DB 4; Length 846;
Best Local Similarity 38.7%; Pred. No. 2.3e-115;
Matches 320; Conservative 128; Mismatches 292; Indels 87; Gaps 21;
QY 4 ICTTRGVSSCCQCLAVSPHCAMCSDDEALPLGSPRCDLKENLLKDKCAEESIEFPSEARV 63
Db 37 VETCSKSKCHTCIOETG-CAMCMQPDFK-GQSRCTYQNTSSL--CPEFAFSPITVBOI 91
QY 64 LEDRPL-----SDKSGDSS-----78
Db 92 LVNNKLTNGYKALAGGGGGMSS 151
QY 79 QVTQVSPQIALRLRPDSKNFSIYQVEDYPVDIYLLMDLSYSKMDLWSTIONLIGTKL 138
Db 152 EIVQIOPQSMRLALRVNEKHNKIKISQAEGYVPDLYLLMDLSKSMEDDKAKLSTLGDKL 211
QY 139 ATQMKLTSLNLRIGGAFVDRKVSPPYMTISPEALENFCYDMKTTCLEMFQYKHVLTITD 198
Db 212 STTKRITNNFHLFGSEFVQKLYMPV-STIPKLEHPC---ENCKAPGYQNMPLNN 266
QY 199 QVTRNEEVKQSVARNRDPAGCFDAIMQATVCDKIGWRNDASHLLVFTTDAKTHAL 258
Db 267 NTESFSEKNAKTVSGNDAPBGGFDALMQALACNSQIGWREQARRLLVFTDAGFHAG 326

QY 259 DGRLAGIOPNDGCHVSDNHHYSASTMDYPSLGMTEKLSOKNINLIFAVTENVNLV 318
Db 327 DKLIGVIAIPNDGECBHPKEGYTHSTLODDYPSISQINKQKADNAINIFAVTASQSLV 386
QY 319 QNYSSELIPGTVGVLSMDSNVNLOLIDVAYKISKVELEVRDLPPELSTLSFNATCLNE 378
Db 387 EKLVEHIGSSAAKLDNDSNVNVELVKEKYREKISSVEMK-DNATGVKITYFPSSCLSN 445
QY 379 VIPGLKSCMGKIGDVTFSFIEAKVRCGPQERE--KSFITKPYGFKDSLIVQTEPDCC 435
Db 446 PEVQTSKCDNLKEGQVVSFTAOILKCPEDPRMTQIHIHPGINEVMOILTMLCSC 505
QY 426 ACQ-----AQAPRSHCNNGNTEFEGYVRCGPGMIGSCGSESEEDYRSDQDECS 489
Db 506 PEENPSSIGYVOQANS--C-SGHGYSKICGICNDOSYFGNKECSAIDLISKFANDPSCR 562
QY 490 EGQPV---CSORGECIGQCVCHS--SDPFGKITGKCECDPFCVRYKGEKSCS--HGOC 542
Db 563 ADSTITDCSGRGHOCVACGCHRPNIELISGKHCECDNFSCERRNQLCSGPDHGT 622
QY 543 SCGDCLCDSDMTGYCNCCTTPTQW-SSNGLLCSGRKCEGSCVCI---QPGSYGDT 598
Db 623 ECGRCKCKPGWTGSCGQESNDTCMPGGGEGELSGHGTCEGVCCTVNDQGRFSGRHC 682
QY 599 ECKPCTPACTFEKKECVCEKFKFDGALHDENTCNRCD---EIESVKELKDTGRDAVN 654
Db 683 ECKPCTSCRCGLKDCVQCCMYKTKGELKNGDCCARNTQYFVGVVEY-ELDET-KDEQM 740
QY 655 CTYKNECDVVRFOYEDSSGKSLIYVEEPECRGPDIIVLLSYGAILLIGLALLI 714
Db 741 CKFDEDCDFEMFYSE--QELHVVYAQENNECPAKFVGLGVIAIVGLAILL 798
QY 715 WKLLTTHDRKEPAKFEERARAKMDTANPLVKEASTFTFNITYRG 761
Db 799 WKLLTTHDRREFARFEKERMANAKMDTGENDPIYKQATSTFKNPMYAG 845

RESULT 14

US-08-789-078-1
Sequence 1, Application US/08789078
Patent No. 5843885
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Stahan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbets, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789,078
FILING DATE: 03-FEB-1997
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/229,513
FILING DATE: 19-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 22833

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: 816)474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORGANISM: Homo sapiens
TISSUE TYPE: tonsil
FEATURE:
NAME/KEY: Region
LOCATION: 1..22
OTHER INFORMATION: /label= "signal sequence"
FEATURE:
NAME/KEY: Region
LOCATION: 445..496
OTHER INFORMATION: /label= "repeat"
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 497..540
OTHER INFORMATION: /label= "repeat"
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 541..581
OTHER INFORMATION: /label= "repeat"
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 582..617
OTHER INFORMATION: /label= "repeat"
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 701..723
OTHER INFORMATION: /label= "trans"
OTHER INFORMATION: /note= "transmembrane domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 724..769
OTHER INFORMATION: /label= "cyto"
OTHER INFORMATION: /note= "cytoplasmic domain"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
TITLE: LFA-1 Amino acid sequence (82) (from human
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 96-96
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
US-08-789-078-1

```

```

Query Match 36.4%; Score 1513.5; DB 2; Length 769;
Best Local Similarity 39.9%; Pred. No. 1,2e-113;
Matches 308; Conservative 133; Mismatches 278; Indels 53; Gaps 21;

```

```

OY 119 DISYKDDLSINQNLGKTLATOMRKLTSNLRIGCAFVDKVPSPYMTISPEALENCPY 178
DB 134 DLSYMLDLDLRNVKRLGDDLALALNEITESGRIGSFVDLTVLPFN-TAPDKLRNCP 192
OY 179 DMKTCILPMFGYKHYLTLLTDQVTRFNEVEKOSVARNADAPGCGDAMOMTVDEKIGW 238
DB 193 NKEEQPEFAFARHVKLTLNNSNOFOTEGKOLISGLNDAPEGGLDAMQVAACPEEIGW 252
OY 239 RNDASHLVFTTDAKTHIALDGRLAGIYVQNDGCGVSDNHSASTTMDYPSGLMTER 298
DB 253 RN-VTRLVFAFDGDFHFAQDGKLGAILLPNDGRCHL-EDNLKYSNEDFPVSQLAHK 310
OY 299 LSKQINILIPAVENVNLKQNTSELIPGTVGVSMSSNVLDLIVAYKIRKAYLE 358
DB 311 LAENNIOPIFAVTSRMVTKYEKLEIPKSAVGELESDSSNVVHLIKNAKLSRYFLD 370
OY 359 VRDLPBELSLFNATCLNNEVIRGLK-----SCMGKIGDGYFSFEAKVRGCPQKE 411
DB 371 HNALPDTLKVYTDSCSN-----GYTHRNQPRGDDCGVQINPITFOVKYVATECIG--E 423
OY 412 KSFTRKPVGFKDSLIVYVTFDCACQAQAPENSHRCNNGTFCGVCRGPGNVLGSGQC 471
DB 424 QSFVIRALGFTDIYVYVLPQCECRDRS-LC-HGKFLGCLRCIDTGYIGKNC 481
OY 472 ECSEEDYRPSQDE--CSFREGQPVCSQREGCLCGQCYCHSDF-GK-ITGKYCCDPS 527
DB 482 ECOTOG-RSSQELBGSCKRDNNSIICSLGDCVCCQCLCHSDYDGLIYQGCCDDIN 540
OY 528 CVRYKEMSGHGO--CSCGDCLCDSDMTGYCNCCTTRTDTCMSNGLLCSGRGCEGS 585
DB 541 CERYNGQVCGGFRGLCFEGKCRCHPGEFSACQCEBRTTEGLNRRVCEGSRGRCNV 600
OY 586 CVCIOPGSYGDCCKOPTCPDACPCKECCCKFDRGALHDENTCNRYCNDEISVREL 645
DB 601 CEC-HSGYQLPLCQCECPGSPCGKYISCAELCKEKP--GKNCSAACPG-----LQL 652
OY 646 KDTGDAVNCYTKNEDDCYVRFQY-YEDSSGSLIYVVEPCPGPOLIVLLSVMAI 704
DB 653 SNNPVKRGTKRDESGCVAATLEQDDMDRYLLIYVESRCCVAGPNAALVGGTVAGI 712
OY 705 LILGLAALLIMKLLITIHDKREFAFEEERARAKMDTANPNLYKATSTFTN 756
DB 713 VLIGILLIYIMKALLHLSDLREYRREKELKSQNN-DNPLFKSATITVMN 763

RESULT 15
US-08-752-633-1
; Sequence 1, Application US/08752633
; Patent No. 5863889
; GENERAL INFORMATION:
; APPLICANT: Benedict, Stephen H.
; APPLICANT: Slahaan, Teruna
; APPLICANT: Chan, Marcia
; APPLICANT: Tibbels, Scott
; TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
; STREET: 1101 Walnut St.
; CITY: Kansas City
; STATE: MO
; COUNTRY: USA
; ZIP: 64106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/752,633
; CLASSIFICATION: 514

```

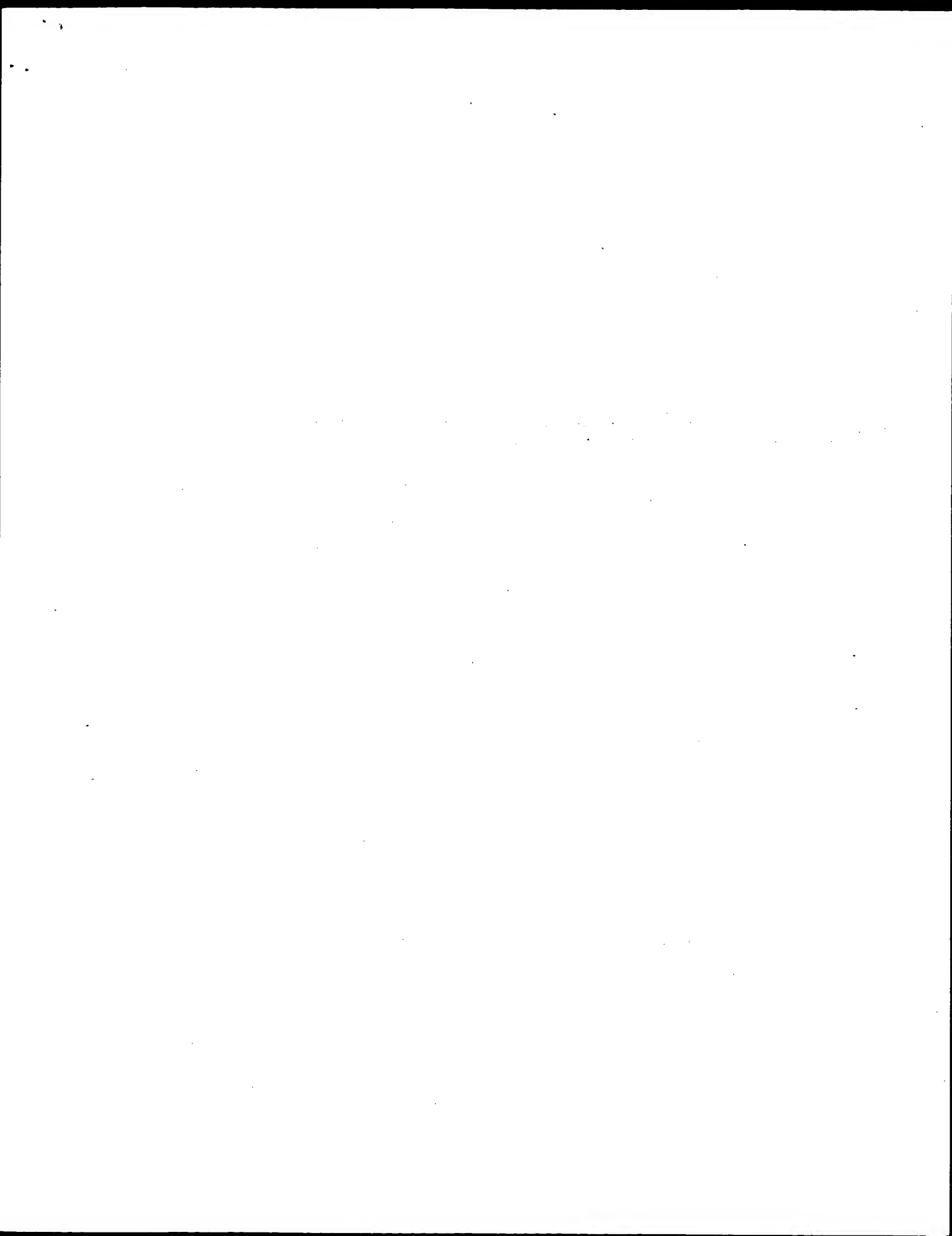

ATTORNEY/AGENT INFORMATION:
 NAME: Collins, John M.
 REGISTRATION NUMBER: 26262
 REFERENCE/DOCKET NUMBER: 22833
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (816)474-9050
 TELEFAX: 434-363
 TELEX: 434-363
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 TOPOLOGY: Both
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 TISSUE TYPE: tonsil
 FEATURE:
 NAME/KEY: Region
 LOCATION: 1..22
 OTHER INFORMATION: /label= signal
 FEATURE:
 NAME/KEY: Region
 LOCATION: 449..496
 OTHER INFORMATION: /label= repeat
 FEATURE:
 NAME/KEY: Region
 LOCATION: 541..581
 OTHER INFORMATION: /note= "cysteine rich repeat"
 FEATURE:
 NAME/KEY: Region
 LOCATION: 497..540
 OTHER INFORMATION: /label= repeat
 FEATURE:
 NAME/KEY: Region
 LOCATION: 582..617
 OTHER INFORMATION: /note= "cysteine rich repeat"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 701..723
 OTHER INFORMATION: /label= trans
 OTHER INFORMATION: /note= "transmembrane domain"
 NAME/KEY: Domain
 LOCATION: 724..769
 OTHER INFORMATION: /label= cyto
 OTHER INFORMATION: /note= "cytoplasmic domain"
 PUBLICATION INFORMATION:
 AUTHORS: Pigott,
 TITLES: LFA-1 Amino acid sequence (B2) (from human
 TITLE: tonsil)
 JOURNAL: The Adhesion Molecule Facts Book
 PAGES: 96-96
 DATE: 1993
 RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
 US-08-752-633-1

Query Match 36.4%; Score 1513.5; DB 2; Length 769;
 Best Local Similarity 39.9%; Pred. No. 1.2e-113;
 Matches 308; Conservative 133; Mismatches 278; Indels 53; Gaps 21;

QY 5 CTTGGVSCQCCIAVSPMCAMCSDALPLGSP-----RCDLKENILKDNCAPESEIEPPV 58
 DB 25 CTKFVSSGRCIEGSGPCTWC--QKLNFTGCGDPDSIRCDTRQLLMRGCAADIMPT 82

QY 59 SEARVLEDRPLSDKSGSDSSQVTVSPPORIALRLRPDSDKSNSTQVROVEDYVPDIYIM 118
 DB 83 SLAEQEDHNGGK-----QLSPQKVTLYLRPGGAALFNVFRRAKGYPLIDLYIM 133
 QY 119 DLSYMKDLDLWSTQNLGTRLATQMRKLTNSRIGFAGVDRPVSVMYISPPALENCY 178
 DB 134 DLSYMDLDRNVKRLGGDLRLALNEIESRIGSGVDTVLPFVN-TRPDRLRNP 192
 QY 179 DMKTCCLPMFGYKHVLTITDOVTRPNEVKKQSVSRNDAPEGGDAIMQATVDEKIGW 238
 DB 193 NKEKCCPPFAFRHVLKLTNNSNQFTEVGKQLISGNLDAPEGGLDMQVAAPPEEIGW 252
 QY 239 RNDASHLVFTTDAKTHALDGRLAGIVQPDGCGHVSQDNHYASATMDVPSGLMTER 298
 DB 253 RN-VTRLVFATDGFHAGGKLGAILTPDGKCHL-EDMLYKRSNEFDPSGQLAHK 310
 QY 299 LSQKNILIFAVTENVMYLYNSSELLPGTTVGVLSMDSNVQLIYDAYKIRKYLE 358
 DB 311 LAENNIPFVATSMVTKETKEIIPKSAVGESEDSNVHLIKNAVKLSRVFLD 370
 QY 359 VRDLPHEELSLFPMATCLNNEVIPLK-----SCMLKIGDIYSEFSEIAKVRGCPPEKE 411
 DB 371 HNALPDTLKVYDSCSN-----GVTRNPDGDCDGVQINVPITFQVKVATECIQ--E 423
 QY 412 KSFTIKPVGFKDSLIVQYTFDCDCAQAAEPNSHRGNNGTEECGVCRCGPGMLGSCQ 471
 DB 424 QSEVIRALGFTDIYVQLPQCCECRCRQSDRS-LC-HKGFLGEGICRQDITGICKNC 481
 QY 472 ECSEDIYPSQDE--CSPREGQVYCSRGKCLGGQCYCHSSD-GR-ITGKICECDFE 527
 DB 482 ECQTQG-RSSQLESGSCRKDNNSIICSLGIDCGVGGCLCHSDVPGKLIQYCECDPTIN 540
 QY 528 CVRYKSGMSGHQ--CSCGDCLDSMTGYCNCCTRTDPCMSNGLCSGRKCEGS 585
 DB 541 CERNYGVCGRPGRLGFCGCKRCHPGEBSACCCERTTEGCLNPRVYECSSGRKRCNV 600
 QY 586 CVCIQPSYSGDTCEKCPDPACTCFKKECYECKFRGALHDENTCNRYCRDEIESYKEL 645
 DB 601 CEC-HSGYQLPLCQCEPCGKISACCLFEKGP--GKNCSAACPG-----LQL 652
 QY 646 KDTGRDANVCYKKNEDCVARFQY-YEDSSKSLIYVEEPECKGPDILVLLSVGAI 704
 DB 653 SNPVKGTCKERSEGCWAAVYTLQDDGMRYLIYVESRECVAGNIAIVGTVAGI 712
 QY 705 LLIGLAILIKLITITHDREFAKFEERARAKMDTANNPLYEASTFTN 756
 DB 713 VLIGILLVIMKALIHLSDLREYRFEKELKSGMNN-DNPLKRSATTVMN 763

Search completed: May 19, 2002, 10:39:46
 Job time: 9115 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 09:42:13 ; Search time 55.89 Seconds
(without alignments)
1310.075 Million cell updates/sec

Title: US-09-673-302a-1

Perfect score: 4154
Sequence: 1 GPNICTTGVSSCCOCLAVS.....NPLEYKKAETSTNTTNGT 762

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pIR.71:*
2: pIR.1:*
3: pIR.3:*
4: pIR.4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4154	100.0	788	2	platelet glycoprot
2	4146	99.8	778	2	platelet glycoprot
3	4119	99.2	788	2	platelet glycoprot
4	4038	97.2	753	2	platelet glycoprot
5	3577	86.1	723	2	integrin beta-3 ch
6	3484.5	83.5	781	2	integrin beta-3 - c
7	3470	83.5	680	2	integrin beta-3 ch
8	3260.5	78.5	788	2	integrin beta-3 su
9	2340	56.3	799	2	integrin beta-5 ch
10	2156	51.9	788	2	integrin beta-6 ch
11	1994.5	46.0	656	2	integrin beta-5 ch
12	1846.5	44.5	803	1	integrin, band 3 p
13	1831	44.1	798	2	fibronectin recept
14	1826.5	44.0	798	2	integrin beta-1* c
15	1821	43.8	798	2	integrin beta-1 ch
16	1818.5	43.8	798	2	integrin beta-1 ch
17	1817	43.7	773	2	beta-1 integrin su
18	1806.5	43.5	799	1	fibronectin recept
19	1801.5	43.4	799	2	integrin beta o1lg
20	1676	40.3	809	2	integrin beta-6 ch
21	1579	38.0	877	2	integrin beta-4 ch
22	1553	37.4	772	2	integrin beta-2 ch
23	1547.5	37.3	846	2	integrin beta-4 ch
24	1520.5	36.6	769	1	leukocyte adhesion
25	1517	36.5	771	2	leukocyte adhesion
26	1515	36.5	770	2	leukocyte adhesion
27	1513.5	36.4	769	1	leukocyte adhesion
28	1506	36.3	806	1	integrin beta-7 ch
29	1502.5	36.2	798	2	integrin beta-7 ch

30	1331	32.0	768	2	B41029	integrin beta-8 ch
31	1326.5	31.9	769	2	A41029	integrin beta-8 ch
32	1184	28.5	964	2	JC5545	integrin beta-4 pr
33	1184	28.5	1875	2	A36429	integrin beta-4 ch
34	1154.5	27.8	1807	2	JC6319	integrin beta-4 ch
35	1145.5	27.6	1748	1	JN0786	integrin beta-4 ch
36	948	22.8	277	2	S14324	platelet glycoprot
37	738	17.8	150	2	I70163	integrin beta-3 su
38	545	13.1	92	2	D37057	beta 8 integrin -
39	419	10.1	194	2	I51310	integrin beta-1 ch
40	272	6.5	95	2	S6396	tenascin precursor
41	268.5	6.5	2019	1	JQ1322	odx protein - fru1
42	265	6.4	2406	2	A54148	tenascin-like prot
43	265	6.4	2515	2	S47008	tenascin-X precurs
44	260.5	6.3	3566	1	A40701	tenascin-C - human
45	259	6.2	2201	2	A32160	

ALIGNMENTS

RESULT 1

177349
platelet glycoprotein IIb precursor - human

N:Alternate names: GPIIb
C:Species: Homo sapiens (man)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 04-Mar-2000
C:Accession: I77349; F56793

R:Frachet, P.; Uzan, G.; Thevenon, D.; Denarier, E.; Prandini, M.H.; Marguerie, G.
Mol. Biol. Rep. 14, 27-33, 1990

A:Title: GPIIb and GPIIb amino acid sequences deduced from human megakaryocyte cDNAs
A:Reference number: 157461; MUID:90265363

A:Accession: I77349
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-788 <FRA>

A:Cross-references: GB:M35999; NID:q183532; PIDN:AAA35927.1; PID:q306795
R:Cailliet, B.; Parmentier, S.; Leung, L.U.; McGregor, J.L.

Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIc*, GPIIa
A:Reference number: A56793; MUID:92061944

A:Accession: F56793
A:Status: preliminary

A:Molecule type: protein
A:Residues: 27-30, 'X', 32-37 <CAP>

A:Experimental source: platelet
C:Superfamily: integrin beta chain; laminin-type EGF-like homology

C:Keywords: glycoprotein; platelet
F:1-21/Domain: signal sequence #status predicted <SIG>

F:123,346,478,782/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 100.0%; Score 4154; DB 2; Length 788;
Best Local Similarity 100.0%; Pred. No. 1.5e-256;

Matches 762; Conservative 100; Mismatches 0; Indels 0; Gaps 0;

QY	1	GPNICTTGVSSCCOCLAVSPMCAMCDEALPLGSPKDLKLNLCNCAEESTFPSE	60
DB	27	GPNICTTGVSSCCOCLAVSPMCAMCDEALPLGSPKDLKLNLCNCAEESTFPSE	86
QY	61	AVLEDRPLSDKSGDSQVQVSPQIALRLRPDSSNFSGIQRVQVEDYVVDIYLLMDL	120
DB	87	AVLEDRPLSDKSGDSQVQVSPQIALRLRPDSSNFSGIQRVQVEDYVVDIYLLMDL	146
QY	121	SYSMKMDLMSIONLGTKLATQWRKLTSLNRIGFGAFVDKPVSPYWTSPPEALENPECYDM	180
DB	147	SYSMKMDLMSIONLGTKLATQWRKLTSLNRIGFGAFVDKPVSPYWTSPPEALENPECYDM	206
QY	181	KTTCPMPGKYKIVLLTQVYTRFNEEVKKQSVSRKRDAPEGGFALIMQATYCDKIGWRN	240
DB	207	KTTCPMPGKYKIVLLTQVYTRFNEEVKKQSVSRKRDAPEGGFALIMQATYCDKIGWRN	266
QY	241	DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHGSDNHNYSATITMDYPSLGIMTEKLS	300

```

Db 267 DASHLVEFTTDAKTHIALDRLAGIVQPNDOGHVSDNHYASSTMDYPSLGLMTEKLS 326
QY 301 OKNINLIFAVTENVVNLKYONSELIPGTGYVLSMDSNVLOLIVAYGIRSKVELEVR 360
Db 327 OKNINLIFAVTENVVNLKYONSELIPGTGYVLSMDSNVLOLIVAYGIRSKVELEVR 386
QY 361 DLPEELSLSFNATCLNNEVTPGLKSCMGLKIGDVSFSEIAKVRGCPQEKESFTIKPYG 420
Db 387 DLPEELSLSFNATCLNNEVTPGLKSCMGLKIGDVSFSEIAKVRGCPQEKESFTIKPYG 446
QY 421 FKDSLIVQVTFDDCCACQAOAEPSNRHRCNNNGTFECGVCRCGPMWLGSCQECSEEDYRP 480
Db 447 FKDSLIVQVTFDDCCACQAOAEPSNRHRCNNNGTFECGVCRCGPMWLGSCQECSEEDYRP 506
QY 481 SQODECSPRREGQPVCSQOREGCLGQCVCVCHSSDPFKITGKYCECDDPSCYRYKGMCSGNG 540
Db 507 SQODECSPRREGQPVCSQOREGCLGQCVCVCHSSDPFKITGKYCECDDPSCYRYKGMCSGNG 566
QY 541 QCSGDCDCLDSMTGYGNCCTRTDTCMSNGLLCSGRGRCGSCVCITQPSYGDTCRK 600
Db 567 QCSGDCDCLDSMTGYGNCCTRTDTCMSNGLLCSGRGRCGSCVCITQPSYGDTCRK 626
QY 601 CPTCPDACTFKKCEVCKKFDGALHDENTCNRYCDELESYKELDKTGKDAVNCTYKNE 660
Db 627 CPTCPDACTFKKCEVCKKFDGALHDENTCNRYCDELESYKELDKTGKDAVNCTYKNE 686
QY 661 DDCVVRFOYVEDSSGKSILYVEEPCPKGPDILVLLSVMGAILLIGLAALLIKMLIT 720
Db 687 DDCVVRFOYVEDSSGKSILYVEEPCPKGPDILVLLSVMGAILLIGLAALLIKMLIT 746
QY 721 IHDRKEPAKFEERARAKWDTANNPPLYKEATSTFTNITYRGT 762
Db 747 IHDRKEPAKFEERARAKWDTANNPPLYKEATSTFTNITYRGT 788

RESULT 2
A:Accession: A60798
A:Molecule type: DNA
A:Residues: 1-778 <ROS>
N:Alternate names: antigen CD61; integrin beta 3
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text-change 23-Jul-1999
R:Rose, J.P.; Bray, P.F.; Gayet, O.; Johnston, G.L.; Cook, R.G.; Jackson, K.W.; Shuman,
A:Title: Cloning of glycoprotein IIa cDNA from human erythroleukemia cells and localization
A:Reference number: A60798; MUID:88294314
A:Accession: A60798
A:Molecule type: mRNA
A:Status: not compared with conceptual translation
A:Residues: 1-778 <ROS>
R:Lanza, F.; Kleffer, N.; Phillips, D.R.; Fitzgerald, L.A.
J. Biol. Chem. 265, 18098-18103, 1990
A:Title: Characterization of the human platelet glycoprotein IIa gene. Comparison with
A:Reference number: A36085; MUID:91009291
A:Accession: A36085
A:Molecule type: DNA
A:Residues: 17-705, 'G', 707-778 <LAN>
A:Cross-references: GB:M57494
R:Hiraiwa, A.; Matsukage, A.; Shiku, H.; Takahashi, T.; Naito, K.; Yamada, K.
Blood 69, 560-564, 1987
A:Title: Purification and partial amino acid sequence of human platelet membrane glycoprotein
A:Reference number: A90731; MUID:87101510
A:Accession: D32528
A:Molecule type: protein
A:Residues: 208-224 <HI2>
A:Accession: C32528
A:Molecule type: protein
A:Residues: 429-433 <HIR>
R:Zimrin, A.B.; Gidwitz, S.; Lord, S.; Schwartz, E.; Bennett, J.S.; White II, G.C.; Pond
J. Biol. Chem. 265, 8590-8595, 1990
A:Title: The genomic organization of platelet glycoprotein IIa.
A:Reference number: A36268; MUID:90256778

```

```

A:Accession: A36268
A:Molecule type: DNA
A:Residues: 18-705, 'G', 707-778 <ZIM>
A:Cross-references: GB:J05427
A:Note: The authors translated the codon GAT for residue 233 as Glu, GAT for residue
C:Genetics:
A:Gene: GDB:117GB3
A:Cross-references: GDB:120013; OMIM:173470
A:Map position: 17q21.32-17q21.32
C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
C:Keywords: alternative splicing; glycoprotein

Query Match 99.8%; Score 4146; DB 2; Length 778;
Best Local Similarity 99.9%; Pred. No. 4.6e-256;
Matches 761; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPNICTTGVSSCOCLAVSPMCAMCSDEALPLGSPRODLKENLKNCAPESEIEPVSE 60
Db 17 GPNICTTGVSSCOCLAVSPMCAMCSDEALPLGSPRODLKENLKNCAPESEIEPVSE 76
QY 61 ARVLEDRPLSKGSDSSQVTVSPORIALRLRPDSSKNFSIOVROYEDYVPDIYLM 120
Db 77 ARVLEDRPLSKGSDSSQVTVSPORIALRLRPDSSKNFSIOVROYEDYVPDIYLM 136
QY 121 SYSKMDLSTIONLGTKLATQMRKLTSMRLRGFAFDKPVSPMYTSPPEALENFCYDM 180
Db 137 SYSKMDLSTIONLGTKLATQMRKLTSMRLRGFAFDKPVSPMYTSPPEALENFCYDM 196
QY 181 KTCLPMGYKXHVLLTDQVTRFNEEVKQSVSRNRDAPRGGPDALMQAVCDEKIGWRN 240
Db 197 KTCLPMGYKXHVLLTDQVTRFNEEVKQSVSRNRDAPRGGPDALMQAVCDEKIGWRN 256
QY 241 DASHLVEFTTDAKTHIALDRLAGIVQPNDOGHVSDNHYASSTMDYPSLGLMTEKLS 300
Db 257 DASHLVEFTTDAKTHIALDRLAGIVQPNDOGHVSDNHYASSTMDYPSLGLMTEKLS 316
QY 301 OKNINLIFAVTENVVNLKYONSELIPGTGYVLSMDSNVLOLIVAYGIRSKVELEVR 360
Db 317 OKNINLIFAVTENVVNLKYONSELIPGTGYVLSMDSNVLOLIVAYGIRSKVELEVR 376
QY 361 DLPEELSLSFNATCLNNEVTPGLKSCMGLKIGDVSFSEIAKVRGCPQEKESFTIKPYG 420
Db 377 DLPEELSLSFNATCLNNEVTPGLKSCMGLKIGDVSFSEIAKVRGCPQEKESFTIKPYG 436
QY 421 FKDSLIVQVTFDDCCACQAOAEPSNRHRCNNNGTFECGVCRCGPMWLGSCQECSEEDYRP 480
Db 437 FKDSLIVQVTFDDCCACQAOAEPSNRHRCNNNGTFECGVCRCGPMWLGSCQECSEEDYRP 496
QY 481 SQODECSPRREGQPVCSQOREGCLGQCVCVCHSSDPFKITGKYCECDDPSCYRYKGMCSGNG 540
Db 497 SQODECSPRREGQPVCSQOREGCLGQCVCVCHSSDPFKITGKYCECDDPSCYRYKGMCSGNG 556
QY 541 QCSGDCDCLDSMTGYGNCCTRTDTCMSNGLLCSGRGRCGSCVCITQPSYGDTCRK 600
Db 557 QCSGDCDCLDSMTGYGNCCTRTDTCMSNGLLCSGRGRCGSCVCITQPSYGDTCRK 616
QY 601 CPTCPDACTFKKCEVCKKFDGALHDENTCNRYCDELESYKELDKTGKDAVNCTYKNE 660
Db 617 CPTCPDACTFKKCEVCKKFDGALHDENTCNRYCDELESYKELDKTGKDAVNCTYKNE 676
QY 661 DDCVVRFOYVEDSSGKSILYVEEPCPKGPDILVLLSVMGAILLIGLAALLIKMLIT 720
Db 677 DDCVVRFOYVEDSSGKSILYVEEPCPKGPDILVLLSVMGAILLIGLAALLIKMLIT 736
QY 721 IHDRKEPAKFEERARAKWDTANNPPLYKEATSTFTNITYRGT 762
Db 737 IHDRKEPAKFEERARAKWDTANNPPLYKEATSTFTNITYRGT 778

RESULT 3
A25547
A:Accession: A25547
A:Molecule type: DNA
A:Residues: 1-778 <ROS>
N:Alternate names: antigen CD61; integrin beta 3
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text-change 23-Jul-1999
R:Rose, J.P.; Bray, P.F.; Gayet, O.; Johnston, G.L.; Cook, R.G.; Jackson, K.W.; Shuman,
A:Title: Cloning of glycoprotein IIa cDNA from human erythroleukemia cells and localization
A:Reference number: A60798; MUID:88294314
A:Accession: A60798
A:Molecule type: mRNA
A:Status: not compared with conceptual translation
A:Residues: 1-778 <ROS>
R:Lanza, F.; Kleffer, N.; Phillips, D.R.; Fitzgerald, L.A.
J. Biol. Chem. 265, 18098-18103, 1990
A:Title: Characterization of the human platelet glycoprotein IIa gene. Comparison with
A:Reference number: A36085; MUID:91009291
A:Accession: A36085
A:Molecule type: DNA
A:Residues: 17-705, 'G', 707-778 <LAN>
A:Cross-references: GB:M57494
R:Hiraiwa, A.; Matsukage, A.; Shiku, H.; Takahashi, T.; Naito, K.; Yamada, K.
Blood 69, 560-564, 1987
A:Title: Purification and partial amino acid sequence of human platelet membrane glycoprotein
A:Reference number: A90731; MUID:87101510
A:Accession: D32528
A:Molecule type: protein
A:Residues: 208-224 <HI2>
A:Accession: C32528
A:Molecule type: protein
A:Residues: 429-433 <HIR>
R:Zimrin, A.B.; Gidwitz, S.; Lord, S.; Schwartz, E.; Bennett, J.S.; White II, G.C.; Pond
J. Biol. Chem. 265, 8590-8595, 1990
A:Title: The genomic organization of platelet glycoprotein IIa.
A:Reference number: A36268; MUID:90256778

```

N:Alternate names: endothelial cell glycoprotein IIA
 C:Species: Homo sapiens (man)
 C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 20-Aug-1999
 C:Accession: A26547; A32532; S65439
 R:Flynn, L.A.; Steinberg, B.; Rall Jr., S.C.; Lo, S.; Phillips, D.R.
 J. Biol. Chem. 262, 3936-3939, 1987
 A:Title: Protein sequence of endothelial glycoprotein IIA derived from a cDNA clone. ID
 A:Reference number: A26547; MUID:87165991
 A:Accession: A26547
 A:Molecule type: mRNA
 A:Residues: 1-788 <FT>
 A:Cross-references: GB:J0703; NID:g183452; PID:AAA52589.1; PID:g306786
 R:Zimlin, A.B.; Elsmann, R.; Vlahire, G.; Schwartz, E.; Bennett, J.S.; Poncz, M.
 J. Clin. Invest. 81, 1470-1475, 1988
 A:Title: Structure of platelet glycoprotein IIA. A common subunit for two different men
 A:Reference number: A32532; MUID:88213696
 A:Accession: A32532
 A:Molecule type: mRNA
 A:Residues: 1-11, 'A', 13-788 <ZIM>
 A:Cross-references: GB:M20311; NID:g190107; PID:AAA60122.1; PID:g190108
 R:Makogonko, E.M.; Yakubenko, V.P.; Ingham, K.C.; Medved, L.V.
 Eur. J. Biochem. 237, 205-211, 1996
 A:Title: Thermal stability of individual domains in platelet glycoprotein IIBIIIA.
 A:Reference number: S65437; MUID:96203926
 A:Accession: S65439
 A:Molecule type: protein
 A:Residues: 27:375;414 <MAK>
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C:Keywords: alternative splicing; cytoskeleton; glycoprotein; transmembrane protein
 F:1-26/Domain: signal sequence status predicted <SIG>
 F:17-788/Product: platelet glycoprotein IIA beta chain #status predicted <MAT>
 F:719-747/Domain: transmembrane #status predicted <TM>
 F:125,346,397,478,585,680,782/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 99.2% Score 4119; DB 2; Length 788;
 Best Local Similarity 99.3% Pred. No. 2, 5e-254;
 Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEPVSE 60
 DB 27 GPNICTTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEPVSE 86
 QY 61 ARVLEDRPLSDKSGSSQVTVSPQRIARLRPPDSKNFSIOVROYEYPDVITYLMDL 120
 DB 87 ARVLEDRPLSDKSGSSQVTVSPQRIARLRPPDSKNFSIOVROYEYPDVITYLMDL 146
 QY 121 SYMKDLDLMSIQNLGTRKLTATQMKRLSNLRIGGAFVDRPVSPYMTISPEALENFCYDM 180
 DB 147 SYMKDLDLMSIQNLGTRKLTATQMKRLSNLRIGGAFVDRPVSPYMTISPEALENFCYDM 206
 QY 181 KTTCLPMFGYKHYLTLDQVTRNEEVKKOSVSRNDAPEGCFDAIMQATVCEKIGMRN 240
 DB 207 KTTCLPMFGYKHYLTLDQVTRNEEVKKOSVSRNDAPEGCFDAIMQATVCEKIGMRN 266
 QY 241 DASHLVLFTTDAKTHIALDGRLAGIVQPNQGCHVSDNHYSASTTMDYPSLGLMTEKLS 300
 DB 267 DASHLVLFTTDAKTHIALDGRLAGIVQPNQGCHVSDNHYSASTTMDYPSLGLMTEKLS 326
 QY 301 OKNNLLEFATENVNLYONVSELIGETVGLSMDSANVLOLIVAYGKIRSKVELEVR 360
 DB 327 OKNNLLEFATENVNLYONVSELIGETVGLSMDSANVLOLIVAYGKIRSKVELEVR 386
 QY 361 DLPELISLSTFNATCLNNEVTPGLSKCMGLKIGTVSFSIEAKVRGPOEKEKSFITKPYG 420
 DB 387 DLPELISLSTFNATCLNNEVTPGLSKCMGLKIGTVSFSIEAKVRGPOEKEKSFITKPYG 446
 QY 421 FKSLIVQVTFDDCCACQAOAEPNSHRCNNGNGTFEGGVCRCGPWGLSGCCESEEDYRP 480
 DB 447 FKSLIVQVTFDDCCACQAOAEPNSHRCNNGNGTFEGGVCRCGPWGLSGCCESEEDYRP 506
 QY 481 SQODEGSPREGOPVCSRGECISGQCVCHSSDREKKTGKCECDPDFSCVYKKGEMCSGHG 540

DB 507 SQODEGSPREGOPVCSRGECISGQCVCHSSDREKKTGKCECDPDFSCVYKKGEMCSGHG 566
 QY 541 QCSGDCDCLDSDMTGYCNCTRTDTQMSNGLLCSGRGKCEGSCVCTOPGSTGTCEK 600
 DB 567 QCSGDCDCLDSDMTGYCNCTRTDTQMSNGLLCSGRGKCEGSCVCTOPGSTGTCEK 626
 QY 601 CPTCPDACTFKKEVECKKPDRCGLHDENCRNRCREISVYKELDTGADVACTYKNE 660
 DB 627 CPTCPDACTFKKEVECKKPDRCGLHDENCRNRCREISVYKELDTGADVACTYKNE 686
 QY 661 DDCVRFQYEDSSGSKITLYVEEPECCKGPDILVLSVGAITLLGLAALLIMKLLIT 720
 DB 687 DDCVRFQYEDSSGSKITLYVEEPECCKGPDILVLSVGAITLLGLAALLIMKLLIT 746
 QY 721 IHRKEFAKFEERARAKMDTANNPLKREATSTNTITTYRGT 762
 DB 747 IHRKEFAKFEERARAKMDTANNPLKREATSTNTITTYRGT 788

RESULT 4
 B36268
 platelet glycoprotein IIA-II - human (fragment)
 N:Alternate names: Integrin beta-3' chain
 C:Species: Homo sapiens (man)
 C:Date: 18-Jan-1991 #sequence_revision 13-Jan-1993 #text_change 20-Aug-1999
 C:Accession: B36268; A33907
 R:Zimlin, A.B.; Gidwitz, S.; Lord, S.; Schwartz, E.; Bennett, J.S.; White II, G.C.; P
 J. Biol. Chem. 265, 8590-8595, 1990
 A:Title: The genomic organization of platelet glycoprotein IIA.
 A:Reference number: A36268; MUID:90256778
 A:Accession: B36268
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-753 <ZIM>
 A:Cross-references: GB:J05427
 A:Note: the authors translated the codon GAT for residue 216 as Glu, GAT for residue
 R:van Kuppevelt, T.H.M.S.M.; Languiño, L.R.; Galile, J.O.; Suzuki, S.; Ruoslahti, E.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5415-5418, 1989
 A:Title: An alternative cytoplasmic domain of the Integrin beta-3 subunit.
 A:Reference number: A33907; MUID:89315807
 A:Accession: A33907
 A:Molecule type: mRNA
 A:Residues: 680-753 <VAN>
 A:Cross-references: GB:M25108; NID:g186502; PID:AAA36121.1; PID:g386833
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C:Keywords: cell adhesion; glycoprotein

Query Match 97.2% Score 4038; DB 2; Length 753;
 Best Local Similarity 99.6% Pred. No. 3, 3e-249;
 Matches 740; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PNICITTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEPVSEA 61
 DB 1 PNICITTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEPVSEA 60
 QY 62 RVLEDRPLSDKSGSSQVTVSPQRIARLRPPDSKNFSIOVROYEYPDVITYLMDL 121
 DB 61 RVLEDRPLSDKSGSSQVTVSPQRIARLRPPDSKNFSIOVROYEYPDVITYLMDL 120
 QY 122 YSMKDLDLMSIQNLGTRKLTATQMKRLSNLRIGGAFVDRPVSPYMTISPEALENFCYDM 181
 DB 121 YSMKDLDLMSIQNLGTRKLTATQMKRLSNLRIGGAFVDRPVSPYMTISPEALENFCYDM 180
 QY 182 TTTCLPMFGYKHYLTLDQVTRNEEVKKOSVSRNDAPEGCFDAIMQATVCEKIGMRN 241
 DB 181 TTTCLPMFGYKHYLTLDQVTRNEEVKKOSVSRNDAPEGCFDAIMQATVCEKIGMRN 240
 QY 242 ASHLVLFTTDAKTHIALDGRLAGIVQPNQGCHVSDNHYSASTTMDYPSLGLMTEKLSQ 301
 DB 241 ASHLVLFTTDAKTHIALDGRLAGIVQPNQGCHVSDNHYSASTTMDYPSLGLMTEKLSQ 300

QY 302 KINILFAVTEVNYVNIYQNSSELLPGTIVGLSMSSNNVLQITVADYAGKIRKVELEVRD 361
 DB 301 KINILFAVTEVNYVNIYQNSSELLPGTIVGLSMSSNNVLQITVADYAGKIRKVELEVRD 360
 QY 362 LPELISLSEFATCLNNEVIFGLKSCGKLIGDTVFSIEAKVNGCQOEKESFTIKPVGR 421
 DB 361 LPELISLSEFATCLNNEVIFGLKSCGKLIGDTVFSIEAKVNGCQOEKESFTIKPVGR 420
 QY 422 KDSLIYOVTFDCCACAOAEPNSHRCNNGNGTFEGGVRCGPGWGLSGCESEEDYRPS 481
 DB 421 KDSLIYOVTFDCCACAOAEPNSHRCNNGNGTFEGGVRCGPGWGLSGCESEEDYRPS 480
 QY 482 QODECSFREGQPVCSQSGECLCGQCVHSSDFGKITGKCECDDEFCVARYKGMCGSHQ 541
 DB 481 QODECSFREGQPVCSQSGECLCGQCVHSSDFGKITGKCECDDEFCVARYKGMCGSHQ 540
 QY 542 CSGGDLCDSDMTGYCNCCTTRDTQMSNGLLCSRGKCEGSCVCIQPGSYGDTCEKC 601
 DB 541 CSGGDLCDSDMTGYCNCCTTRDTQMSNGLLCSRGKCEGSCVCIQPGSYGDTCEKC 600
 QY 602 PTPDPACTFEKCEKCEKFPDRCALHDEMTNRCYCRDEISVVELKDTGKDAVNCYTKNED 661
 DB 601 PTPDPACTFEKCEKCEKFPDRCALHDEMTNRCYCRDEISVVELKDTGKDAVNCYTKNED 660
 QY 662 DCVAREQYEDSSGKSLVVEEPCPKGPDILVLLSYMGAAILLIGLALLIKLLITI 721
 DB 661 DCVAREQYEDSSGKSLVVEEPCPKGPDILVLLSYMGAAILLIGLALLIKLLITI 720
 QY 722 HDKKEFAKFEERAKAKMTANN 744
 DB 721 HDKKEFAKFEERAKAKMTANN 743

RESULT 5

PN0509

Integrin beta-3 chain - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999

C:Accession: PN0509

R:Clontat, A.M.; Rosa, J.P.; Letourneur, F.; Poncez, M.; Rifaat, S.

Biochem. Biophys. Res. Commun. 193, 771-778, 1993

A:Title: A comparative analysis of cDNA-derived sequences for rat and mouse beta-3 integrin

A:Reference number: PN0509; MUID:93290675

A:Accession: PN0509

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-723 <CITE>

C:Superfamily: Integrin beta chain; laminin-type EGF-like homology

C:Keywords: cell adhesion; duplication; heterodimer; membrane protein

Query Match

Best Local Similarity 86.1%; Score 3577; DB 2; Length 723;

Matches 650; Conservative 42; Mismatches 31; Indels 0; Gaps 0;

QY 40 LKLNKLNCAPESTIEFPVSEARVIEDRLSDKSGSSQVTVSPQRIALRLRDDS 99
 DB 1 LKLNKLNKCPESIEFPVSEARVIEDRLSDKSGSSQVTVSPQRIALRLRDDS 99
 QY 100 FSIQROVEDYVDIYLLMDLSYMKDLMSTONIGTKLATQMKLTNLRIGFCAFVDK 159
 DB 61 FSIQROVEDYVDIYLLMDLSYMKDLMSTONIGTKLATQMKLTNLRIGFCAFVDK 120
 QY 160 PVSPLYMISPPALENPCVDMKTCLPMFGYKHLVTLFDQVYRFEVEKKOSVSNRDP 219
 DB 121 PVSPLYMISPPALENPCVDMKTCLPMFGYKHLVTLFDQVYRFEVEKKOSVSNRDP 180
 QY 220 ECGPAIMQATVCEKIGKRNDSHLVFTTDAKTHIALDGLAGIYVQNDGQCHVSDN 279
 DB 181 ECGPAIMQATVCEKIGKRNDSHLVFTTDAKTHIALDGLAGIYVQNDGQCHVSDN 240
 QY 280 HYSASTMDYPSLGMTKLSOKNINLFAVTEVNYVNIYQNSSELLPGTIVGLSMSSNN 339
 DB 280 HYSASTMDYPSLGMTKLSOKNINLFAVTEVNYVNIYQNSSELLPGTIVGLSMSSNN 339

DB 241 HYSASTMDYPSLGMTKLSOKNINLFAVTEVNYVNIYQNSSELLPGTIVGLSMSSNN 300
 QY 340 VQLIYVDATGKIRKVELEVRDLPELISLSEFATCLNNEVIFGLKSCGKLIGDTVFSI 399
 DB 301 VQLIYVDATGKIRKVELEVRDLPELISLSEFATCLNNEVIFGLKSCGKLIGDTVFSI 360
 QY 400 EAKVNGCQOEKESFTIKPVGRKDSLIYOVTFDCCACAOAEPNSHRCNNGNGTFEGGV 459
 DB 361 EAKVNGCQOEKESFTIKPVGRKDSLIYOVTFDCCACAOAEPNSHRCNNGNGTFEGGV 420
 QY 460 CRGPGWGLSGCESEEDYRPSQODECSFREGQPVCSQSGECLCGQCVHSSDFGKITGK 519
 DB 421 CRGPGWGLSGCESEEDYRPSQODECSFREGQPVCSQSGECLCGQCVHSSDFGKITGK 480
 QY 520 YECBDFSCVYRKGMCGSHQSCGDLCDSDMTGYCNCCTTRDTQMSNGLLCSRG 579
 DB 481 YECBDFSCVYRKGMCGSHQSCGDLCDSDMTGYCNCCTTRDTQMSNGLLCSRG 540
 QY 580 KCEGSCVCIQPGSYGDTCEKCPTCPDPACTFEKCEKCEKFPDRCALHDEMTNRCYCRDEI 639
 DB 541 KCEGSCVCIQPGSYGDTCEKCPTCPDPACTFEKCEKCEKFPDRCALHDEMTNRCYCRDEI 600
 QY 640 ESKVRLKDTGKDAVNCYTKNEDCVAREQYEDSSGKSLVVEEPCPKGPDILVLLS 699
 DB 601 ESKVRLKDTGKDAVNCYTKNEDCVAREQYEDSSGKSLVVEEPCPKGPDILVLLS 660
 QY 700 VMGAILLIGLALLIKLLITIHDKKEFAKFEERAKAKMTANNPLYKATSTFTNITY 759
 DB 661 VMGAILLIGLALLIKLLITIHDKKEFAKFEERAKAKMTANNPLYKATSTFTNITY 720
 QY 760 RGT 762
 DB 721 RGT 723

RESULT 6

S43534

Integrin beta3 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 07-Sep-1994 #sequence_revision 26-May-1995 #text_change 20-Aug-1999

C:Accession: S43534; A49312

R:Minura, H.; Ross, F.P.; Chiba, M.; Teitelbaum, S.L.

Endocrinology 134, 1061-1066, 1994

A:Title: 1,25-dihydroxyvitamin D(3) transcriptionally activates the beta(3)-integrin

A:Reference number: S43534; MUID:94164000

A:Accession: S43534

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-781 <MIM>

A:Cross-references: EMBL:X72378; NID:9474038; PIDN:CA51069.1; PID:9474039

A:Note: the authors translated the codon TTT for residue 174 as GLY, GGC for residue

R:Caio, X.; Ross, F.P.; Zhang, L.; Macdonald, P.N.; Chappel, J.; Teitelbaum, S.L.

J. Biol. Chem. 268, 27371-27380, 1993

A:Title: Cloning of the promoter for the avian integrin beta-3 subunit gene and its r

A:Reference number: A49312; MUID:94086557

A:Accession: A49312

A:Molecule type: DNA

A:Residues: 1-26 <CAO>

A:Cross-references: GB:X75348; NID:9441308; PIDN:CA53095.1; PID:9452854

C:Superfamily: Integrin beta chain; laminin-type EGF-like homology

Query Match

Best Local Similarity 83.0%; Score 3484.5; DB 2; Length 781;

Matches 632; Conservative 58; Mismatches 70; Indels 1; Gaps 1;

QY 1 GPNICTTGVSSCQCLAVSPMCASDEALPGSPRCLEKLNKLNCAPESTIEFPVE 60
 DB 20 GSNICATRGVTSCKQCLAVSPMCASAEVVAQSTFRCDLFAULLNGCGRPTEPRSS 79
 QY 61 ARVLEDRPLSDKSGSSQVTVSPQRIALRLRDDSKNFSIOVVEDYVDIYLLMDL 120
 DB 80 ITVLEERPLSDKSGSGST-TTQMSPORIQLNLRPDDSSOMERHVRVROVEDYVDIYLLMDL 138


```

QY 120 LVSMSKDDLSIONLGTATOMRKLTSLNRIGFAPVDKPSVYWIYSPEALNDPCYD 179
D 144 LVSMSKDDLSIONLGTATOMRKLTSLNRIGFAPVDKPSVYWIYSPEALNDPCYD 203
QY 180 MKTTCLPMFGYKHYLTLTDVTRFNEBKQSVSRNDAPEGGFDAIMQATVDEKIGWR 239
D 204 FNFECMPTFGYKHYLTLTEEVLRNNEVQKQSVSRNDAPEGGFDAIMQATVDEKIGWR 263
QY 240 NDASHLLVFTTDAKTHIALDGRLAGIYQPDGCGHSGDNHYSASTTMDYPSLGLMTEKL 299
D 264 NESSHLLVFTTDAKTHIALDGRLAGIYQPDGCGHSLDNDYAASTLDPSTGLMTEKL 323
QY 300 SOKINILFVTEVNVVLYONYSELIPGTGVGLVLSMDSSVNLDIYADAKIKSKYELVY 359
D 324 SOKINILFVTEVNVVLYONYSELIPGTGVGLVLSMDSSVNLDIYADAKIKSKYELVY 383
QY 360 RDLPEELSLSFNATCLNNEVYIPGLKSCMGKIGDVTYSFSLEAKYKGCPOEKESFTIKPY 419
D 384 RDLPEELSLSFNATCLNNEVYIPGLKSCMGKIGDVTYSFSLEAKYKGCPOEKESFTIKPY 443
QY 420 GFNDLSLYVTPDCCACQAAEPNSHRCNNGCTFECVYRCRGPGMLSSQCCSEEDYR 479
D 444 GFNDLSLYVTPDCCACQAAEPNSHRCNNGCTFECVYRCRGPGMLSSQCCSEEDYR 503
QY 480 PSQODECSPEGQPVCSQREGLCGQCVCHSDPFGKITGKYECDDPFCVRYKGMCSGH 539
D 504 PSQODECSPEGQPVCSQREGLCGQCVCHSDPFGKITGKYECDDPFCVRYKGMCSGH 563
QY 540 GQCSGDCCLDSDMTGYCTCTRTDTCMSSNGLCSGKRCGSCGVCVLIQESYDCE 599
D 564 GQCSGDCCLDSDMTGYCTCTRTDTCMSSNGLCSGKRCGSCGVCVLIQESYDCE 623
QY 600 KCPTCPDCTFKECEYCKFKEDGALHDETCNRYCRDETESYKELDKDGAIVNTYKN 659
D 624 KCPTCPDCTFKECEYCKFKEDGALHDETCNRYCRDETESYKELDKDGAIVNTYKN 683
QY 660 EDDCVNFFQYEDSSGKSLIYVEEPCPKGPDILVLLSVMGAILLIGLALLIKLLI 719
D 684 EDDCVNFFQYEDSSGKSLIYVEEPCPKGPDILVLLSVMGAILLIGLALLIKLLI 743
QY 720 TIHDRKEFAFEERAKAKWDTANPNLYKEATSTFNITTYRG 761
D 744 TIHDRKEFAFEERAKAKWDTANPNLYKEATSTFNITTYRG 785

RESULT 9
A38308
Integrin beta-5 chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 20-Aug-1999
C:Accession: A38308; A35775; S12534; S11708
R:McLaren, J.W.; Vestal, D.J.; Chersesh, D.A.; Bodary, S.C.
J. Biol. Chem. 265, 17126-17131, 1990
A:Title: CDNA sequence of the human Integrin beta-5 subunit.
A:Reference number: A38308; MUID:91009141
A:Accession: A38308
A:Molecule type: mRNA
A:Residues: 1-799 <MCU>
A:Cross-references: GB:J05633; NID:9186504; PIDN:AAA59183.1; PID:9186505
A:Note: Parts of this sequence, including the amino end of the mature protein, were conf
R:Suzuki, S.; Huang, Z.S.; Tanihara, H.
Proc. Natl. Acad. Sci. U.S.A. 87, 5354-5358, 1990
A:Title: Cloning of an integrin beta subunit exhibiting high homology with integrin beta
A:Reference number: A35775; MUID:90319111
A:Accession: A35775
A:Molecule type: mRNA
A:Residues: 1-192, 'A', 194-644, 'L', 646-789, 793-799 <SUZ>
A:Cross-references: GB:M35011; NID:9184524; PIDN:AAA52707.1; PID:9306894
R:Ramswamy, H.; Hemler, M.E.
EMBO J. 9, 1561-1568, 1990
A:Title: Cloning, primary structure and properties of a novel human integrin beta subunit
A:Reference number: S12534; MUID:90228356

```

```

A:Accession: S12534
A:Molecule type: mRNA
A:Residues: 1-644, 'L', 646-799 <RAM>
A:Cross-references: EMBL:X53002; NID:933952; PIDN:CAA37188.1; PID:933953
C:Genetics:
A:Gene: GDB:ITGB5
A:Cross-references: GDB:128005; OMIM:147561
A:Map position: 17q11-17qter
C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
C:Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-799/Product: integrin beta-5 chain #status experimental <MAT>
F:25-719/Domain: extracellular #status predicted <EXT>
F:463-513/Domain: laminin-type EGF-like homology <LEG>
F:720-742/Domain: transmembrane #status predicted <TM>
F:743-799/Domain: intracellular #status predicted <INT>
F:347,460,477,505,552,586,654,705/Binding site: carbohydrate (Asn) (covalent) #status

```

Query Match	Score	DB 2	Length	799
Best Local Similarity	55.5%	Pred. No. 3.8e-141		
Matches 433	Conservative 118	Mismatches 205	Indels 24	Gaps 11

```

QY 1 GNICTTRGVSSQOCLAVSPWCAMCSDEALPLGSP-----RCDLKENTLKDNCAPESIE 55
D 24 GNICTSGATSCBECCLLIHPKAWCSKE--DFGSPRISITSRCDLRANLVKNGCGE-IE 80
QY 56 FVYSARVLEDRPLSDKSGSDSS-QYVQSFQRIALRLRPDSKNFSIQVQVQVEDYPIYDI 114
D 81 SPASSFHYLRSLPLSSKSSGAGMDVIQMTQDEIAVLRNPDGKTFYFOLQVQVQVEDYPIYDI 140
QY 115 YIIMDLVSMSKDDLSIONLGTATOMRKLTSLNRIGFAPVDKPSVYWIYSPEALNDPCYD 174
D 141 YIIMDLVSMSKDDLSIONLGTATOMRKLTSLNRIGFAPVDKPSVYWIYSPEALNDPCYD 199
QY 175 NFC--YDMKTCLPMFGYKHYLTLTDVTRFNEBKQSVSRNDAPEGGFDAIMQATVY 232
D 200 NFCIYKLFPPNCVPSFGFRHLPLTRDVSFNEEVRKQSVSRNDAPEGGFDAIMQATVY 259
QY 223 DEKIRMDASHLLVFTTDAKTHIALDGRLAGIYQPDGCGHSGDNHYSASTTMDYPSL 292
D 260 DEKIRMDASHLLVFTTDAKTHIALDGRLAGIYQPDGCGHSGDNHYSASTTMDYPSL 319
QY 293 GIMTEKLSQKNINLFEATENVVNLVONYSELIPGTGVGLVLSMDSSVNLDIYADAKIK 352
D 320 GIMTEKLSQKNINLFEATENVVNLVONYSELIPGTGVGLVLSMDSSVNLDIYADAKIK 379
QY 353 SKVELEVRDLPBELSLSFNATCLNNEVYIPGLKSCMGKIGDVTYSFSLEAKYKGCPOEKE 411
D 380 SKVELEVRDLPBELSLSFNATCLNNEVYIPGLKSCMGKIGDVTYSFSLEAKYKGCPOEKE 439
QY 412 KSFTKLPVGFNDLSLYVTPDCCACQAAEPNSHRCNNGCTFECVYRCRGPGMLSSQCCSE 471
D 440 KSFTKLPVGFNDLSLYVTPDCCACQAAEPNSHRCNNGCTFECVYRCRGPGMLSSQCCSE 498
QY 472 ECSEEDYRPSQODECSPEGQPVCSQREGLCGQCVCHSDPFGKITGKYECDDPFCVRYKGM 531
D 500 ECSEEDYRPSQODECSPEGQPVCSQREGLCGQCVCHSDPFGKITGKYECDDPFCVRYKGM 558
QY 532 KGMCSGHGQCSGDCCLDSDMTGYCTCTRTDTCMSSNGLCSGKRCGSCGVCVLIQESYDCE 591
D 559 KGMCSGHGQCSGDCCLDSDMTGYCTCTRTDTCMSSNGLCSGKRCGSCGVCVLIQESYDCE 618
QY 592 GSYGTCTCKCTCPDCTFKECEYCKFKEDGALHDETCNRYCRDETESYKELDKDGAIVN 650
D 619 GSYGTCTCKCTCPDCTFKECEYCKFKEDGALHDETCNRYCRDETESYKELDKDGAIVN 677
QY 651 DAVNCTTYNEDDCVVRFPQYEDSSGKSLIYVEEPCPKGPDILVLLSVMGAILLIGLALLIGLA 710
D 678 DAVNCTTYNEDDCVVRFPQYEDSSGKSLIYVEEPCPKGPDILVLLSVMGAILLIGLALLIGLA 737
QY 711 ALLIMKLLITTIHDRKEFAFEERAKAKWDTANPNLYKEATSTFNITTYRG 762
D 738 ALLIMKLLITTIHDRKEFAFEERAKAKWDTANPNLYKEATSTFNITTYRG 785

```

Db 738 LLAIMKLLVTHIDRREFAFQSESRNARIEAMSNPLTKRPSTHYDFTFNKFNKSYNCT 797

RESULT 10
 A37057
 Integrin beta-6 chain - human
 C:Species: Homo sapiens (man)
 C:Date: 15-Feb-1991 #sequence_revision 13-Sep-1991 #text_change 19-Jan-2001
 C:Accession: A37057; 169201
 R:Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pytela, R.
 J. Biol. Chem. 265, 11502-11507, 1990
 A:Title: Complete amino acid sequence of a novel integrin beta subunit (beta6) identified
 A:Reference number: A37057; MUID:90307659
 A:Accession: A37057
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-788 <SHE>
 A:Cross-references: GB:M35198; GB:J05522; NID:g186506; PID:AAA36122.1; PID:g186507
 R:Jiang, W.M.; Jenkins, D.; Yuan, Q.; Leung, E.; Choo, K.H.; Watson, J.D.; Kristensen, G.
 Int. Immunol. 4, 1031-1040, 1992
 A:Title: The gene organization of the human beta 7 subunit, the common beta subunit of t
 A:Reference number: 154749; MUID:93002753
 A:Accession: 169201
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 116-157, 'R', 159-197 <JTA>
 A:Cross-references: GB:S45380; NID:9257588; PID:AA823690.1; PID:9257589
 C:Genetics:
 A:Gene: GDB:ITGB6
 A:Cross-references: GDB:131392; OMIM:147558
 A:Map position: 2pter-qter
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C:Keywords: blocked amino end; cell adhesion; cytoskeleton; glycoprotein; lipoprotein; m
 F:708-730/Domain: transmembrane #status predicted <RMB>
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F:7/Binding site: palmitate (Cys) (covalent) #status predicted
 F:16,48,97,260,387,396,463,471/Binding site: carboxyrate (Asn) (covalent) #status predi

Query Match 51.9%; Score 2156; DB 2; Length 788;
 Best local similarity 50.6%; Pred. No. 1.9e-129;
 Matches 389; Conservative 133; Mismatches 231; Indels 16; Gaps 8;

Qy 1 GPNICITR-----GVSSCOQCLAVSPWCAMCSDAL--PLG-SPRCDLKENLKNCADE 52
 Db 14 GRNDSRTMLCLGAGTCECDCLIGPOCAMCAQENFTHPSGVCERCDTPANLLACQLN 73
 Qy 53 STEEPSEARVLEDRPLSKSGSDSQVQVQSPQIALRLPRDDSKNFSIQVQVQVEDPV 112
 Db 74 FLENPSOYEILKNKPLSGROKNSSDIYQIAQSILTLKRGGAQTLOVHVRQTEDEPV 133
 Qy 113 DIYVLDLSYMKDLSIQLSTKATQMRKLTSLRIGFAGFYQKPYSPMYTSPPEA 172
 Db 134 DLYVLDLSAMDDLTNTIKELSGSKSKMSKLTSLNFRIGFSGFYKPYSPVVK--TTPE 192
 Qy 173 LENPCDMMKTTCLPMFGYHVLTLDOVTRFNEEYKQKOSVSRNRDAPEGCFDAIMQATVC 232
 Db 193 IANPCSSIFYCLPFGFPHILPLTDAERFNEIYKQKISANILITPEGGFPAIMQAAVC 252
 Qy 223 DEKIGRNDASHLLVTTDAKTHIALDGLAGIYQVQNDQCVGSDNHNASATMDYPSL 292
 Db 253 KEKIGRNDASHLLVTTDAKTHIALDGLAGIYQVQNDQCVGSDNHNASATMDYPSL 312
 Qy 293 GLETKRNDASHLLVTTDAKTHIALDGLAGIYQVQNDQCVGSDNHNASATMDYPSL 352
 Db 313 GGLIDKLVQNNVLLFAVYQEQVHLENAKILPATVGLLOKQSDNINILQILIAVEELR 372
 Qy 353 SKVELVRLPEELSLFNATCLNNEVYIGLSCMGKLTIGTVYSIEAKVNGCQOEK 412
 Db 373 SEVELVRLPEELSLFNATCLNNEVYIGLSCMGKLTIGTVYSIEAKVNGCQOEK 431
 Qy 413 SFIRKVGKDSLIYQVQNDQCVGSDNHNASATMDYPSL 472

Db 432 HIITKPVGLGALBELVSPENCOCQKEVEVNSSKCHHGNSFEGCVCACHPGMRCE 491

Qy 473 CSEEDYRPSQODECSPRGQPVCSQREGCLGQCVCHSDFGKITGKYCECDDEFCVRYK 532
 Db 492 CGED---MLSTDSCKEAPDRHPSGSGRGDCYGGCICHLSPYGNLGYPCOCODNFSVYRK 548

Qy 533 GEMSGHGGCSCGDCDLSMTGYCNCYTRTPDMSSNGLGSGRGKCEGSCVC10PG 592
 Db 549 GLLCGNGDCCGRCVCSGWTGECYCTSTDCVSDGCLGSRDVCYCKCVCCTNPG 608

Qy 593 SYGDPCEKCPCPACPFKEKCEVCKKPRDGLADEMTCCNRYCDELTESYELKDCD 651
 Db 609 ASGPCEKCPCPGCPDPCNSKSCICDHLSAAGAGE--CYDKCLAGATVSEEDFSKDG 666

Qy 652 AVNCTYKNEDDCVARFOYEDSSGKSLVYVEEPCRPGLVLLSYMGAILLGLAA 711
 Db 667 SVSCSLGECNCLITFLTITDNEKTIHSHNEKDCPRPMIPLMGLVSLATLILVVL 726

Qy 712 LLMKLLITTHIDRKEFAKFEERARAKWDTANPLYEANSTFNNIYR 760
 Db 727 LCIWLLVSPFDRKEVAKFEARAKAKWQGTNPILYKGSSTFNNIYR 775

RESULT 11
 J02005
 Integrin beta-5 chain - baboon
 C:Species: Papio sp. (baboon)
 C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 21-Jan-2000
 C:Accession: J02005
 R:Shoji, M.; Hayzer, D.J.; Kim, T.M.; Runge, M.S.; Hanson, S.R.
 Gene 133, 307-308, 1993
 A:Title: Human and baboon integrin beta 5 subunit-encoding mRNAs have alternative pol
 A:Reference number: J02005; MUID:94040831
 A:Accession: J02005
 A:Molecule type: mRNA
 A:Residues: 1-656 <SHO>
 A:Note: The authors translated the codon AGA for residue 454 as Lys, CNA for residue
 C:Superfamily: This protein is a predominant subunit for the vitronectin receptor in babo
 C:Keywords: cell adhesion
 F:320-370/Domain: laminin-type EGF-like homology <LEG>

Query Match 48.0%; Score 1994.5; DB 2; Length 656;
 Best local similarity 55.6%; Pred. No. 2.9e-119;
 Matches 365; Conservative 102; Mismatches 175; Indels 15; Gaps 7;

Qy 118 MDLSYMKDLSIQLSTKATQMRKLTSLRIGFAGFYQKPYSPMYTSPPEALENPC 177
 Db 1 MDLSYMKDLDITIRNLGTKLAEEMRKLTSLNFRIGFSGFYVDKDISPSY--VAPRYQTNP 59

Qy 178 --YDKRTCLPMFGYHVLTLDOVTRFNEEYKQKOSVSRNRDAPEGCFDAIMQATVCDEK 235
 Db 60 IGYKLFPCVCSFGRHILPLTDVDSNENYKQKORSNRDAPEGCFDAIMQATVCDEK 119

Qy 236 IGRNDASHLLVTTDAKTHIALDGLAGIYQVQNDQCVGSDNHNASATMDYPSLGLM 295
 Db 120 IGRNDASHLLVTTDDVPHALDGLAGIYQVQNDQCVGSDNHNASATMDYPSLGLM 179

Qy 296 TEKSQKINILFAVTEVNVLYQNSSELIGTVYGLSDSSNVLDLYDAYKITSKY 355
 Db 180 GEKLAENINILFAVTKNHYLYKNFALIPGTVVEILDGSKNIDLIINAVYSISKV 239

Qy 356 ELEVDRLPEELSLFNATCLNNEVYIGLSCMGKLTIGTVYSIEAKVNGCQOEK 414
 Db 240 ELVWDQEDENLFTYATCQGVSPYQORCEGLKIDTASFEVSVARCSPSNTHHF 299

Qy 415 TIKPVGKDSLIYQVQNDQCVGSDNHNASATMDYPSL 474
 Db 300 ALOPVGKDSLEVGTVYCTCGCGSVGLEPNSARC--SETGYVGLCELSPGYGLTRCEQ 358

Qy 475 BEDRPQSDSCSPREGQPVCSQREGCLGQCVCHSDFGKITGKYCECDDEFCVRYK 534

Db 359 DGENH5YQNLCDRTBEGKPLCSGRGDCSCNOCSEFSEFGKIVGPCEDNFCANRKGV 418
 QY 535 MCSGHCQSCGDCDLCSDMTGYCNCCTRTDTCMSNGLLCSGRKCEGSCVCIOPGSY 594
 Db 419 LCGSHGCHGCECKCHAGYIGDNMCNSTDSTCRGRDQGLCSRGHCLCGCCCTEPGAF 478
 QY 595 GDTCEKCPDCACTEKECECKKFDRCALHDEMTNCRKDEIES-VKELDTGSKDAV 653
 Db 479 GEMCEKCPDCACTEKECECKKFDRCALHDEMTNCRKDEIES-VKELDTGSKDAV 537
 QY 654 NCYTKNDDCVNRFQYEDSSGKILYVVEPECPKPDILVLLSVMAGILLIGLALL 713
 Db 538 LCFKTKAKDCVMMFTYELPSGKSNLTLYLREPECGMTPNMTLLLVAGSILLVGLALLA 597
 QY 714 IMKLLITHDRKEFAKEEERAKMDTANNPLYKEATSTFT-----NITRYGT 762
 Db 598 IMKLLVTHDRREPAKQSRERARERAMSNPLYKRPISHTYDTFNNKNSYNT 654

RESULT 12

IUCH3
 Integrin, band 3 precursor - chicken
 N:Alternate names: CSAR antigen; JG22 antigen; RGD-receptor
 C:Species: Gallus gallus (chicken)
 C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 23-Jul-1999
 C:Accession: A23947
 R:Rankun, J.W.; Desimone, D.W.; Fonda, D.; Patel, R.S.; Buck, C.; Horwitz, A.F.; Hynes, Cell 46, 271-282, 1986
 A:Title: Structure of integrin, a glycoprotein involved in the transmembrane linkage bet
 A:Reference number: A23947; MUID:86245073
 A:Accession: A23947
 A:Molecule type: mRNA
 A:Residues: 1-803 <TAM>
 A:Cross-references: GB:M14049; NID:9212213; PIDN:AAA48926.1; PID:9212214
 A:Experimental source: embryonic fibroblasts
 A:Note: the amino end of the mature protein is blocked
 C:Comment: Integrin, an integral plasma membrane complex of three (two in mammalian cell
 C:Comment: This transmembrane complex may be the target of oncogenic transformation that
 C:Comment: The cytoplasmic domain of this subunit contains a potential tyrosine-kinase pho
 C:Comment: The extracellular domain of this protein, like many membrane receptors, conta
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C:Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein
 F:1-24/Domains: signal sequence #status predicted <SIG>
 F:25-803/Product: Integrin, band 3 #status predicted <MAM>
 F:25-733/Domains: extracellular #status predicted <EXT>
 F:467-654/Region: cysteine-rich
 F:482-564/Domains: duplication
 F:734-756/Domains: transmembrane #status predicted <MEM>
 F:757-803/Domains: intracellular #status predicted <INT>
 F:216/273,367,410,421,433,445,486,525,589,624,674/Binding site: carbohydrate (Asn) (cova
 F:788/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 44.5%; Score 1846.5; DB 1: Length 803;
 Best Local Similarity 45.1%; Pred. No. 9,4e-110;
 Matches 353; Conservative 136; Mismatches 266; Indels 27; Gaps 15;

QY 1 GPNICTRGVSSCOCCIAVSPKAWC-SDEALPLGSP---RDLKENLKLKDNCAPESTFE 56
 Db 27 GGSDCIRAKNAKSCGECIOAGPNCGWCKKIDFLOEGEPTAKRCDLALAKSKGCEPDIEDN 86
 QY 57 PSEARVLEDRPLSDKSGDSO-----VTVSPQRIALRLRPDSKSNFSIOVRQVEDPY 111
 Db 87 PRGSKRVLEDRVETNRKIGAAEKLEPAITQIQPKLVIQLRNGEPTFSIMKRRADXP 146
 QY 112 VDIYIYLDLSTYSKMDLMSIQNLGTALATOMKRLTNSLRIGAFYVDKVPYSPTWSPPE 171
 Db 147 IDLYIYLDLSTYSKMDLMSIQNLGTALATOMKRLTNSLRIGAFYVDKVPYSPTWSPPE 206
 QY 172 ALNPPCDMKTGLPMGKYHYVLTLDQVTRNEEYKOSYSNRRAPGSPAIQAVY 231
 Db 207 LRRPCTG-DONCTSPSPSYKNVLSITSEGNKEVELGKOHISGNDLSPGSGDALIMQAV 264

QY 232 CDKTIQWRNDASHLLVFTTDAKTHIALDGLAGIVQPDGQCHGSDNHNYSATTTMDYPS 291
 Db 265 CDDQJGWRN-VYRLVLEADGFEAGDGKLGIGIVLPDQKCHL-ENNMVIMSHYIDYPS 322
 QY 292 LGLMTKLSQKINILIFAVTEVNVVLXONSLITGTVGLSNDSSVVLQIYDAVKI 351
 Db 323 IAHVQKLSENNIOITFAVTEEVAVYKELNLPKSAVGLTSSNNSVVLQIYDAVNSL 382
 QY 352 RSKVELEVNDLPEETLSLSEFNATCLN-NEVITGLKSCGMLKIGTVFSIAKRYGCPQE 409
 Db 383 SBEVILLENKLPKREYITISKYCKRGNVDQEDGKCKNISIGDEVREINVTANNECKK 442
 QY 410 -KEKSEFTIKPVGKDSLIVQYTFDCAQQAQAEPSNRHRCNNNGTEFCGVCRCGPGWLG 468
 Db 443 GNETIKIKPLGFTBEVEIHIOFDICLQSEGEPSNBPACHDGNTEFCGACRCNEGRIG 502
 QY 469 SCQCESEDEYRSQODE-CSPREQPVQSOEGELCGGCVQHSBPFGK--ITGKYCEGD 525
 Db 503 RLCESTIDEVNSMEDDAYCRRNSTEICSNNGECIOGCVCKKRENTNEVYSGKICECDN 562
 QY 526 FSCVRYKGMCSGHCQSCGDCDLCSDMTGYCNCCTRTDTCMSNGLLCSGRKCEGCS 585
 Db 563 FNCDSRNGLLCGGNGICRCVCEPNTFGSACDCLDTPCOMANGQICNGRGTCEGT 622
 QY 586 CYCIOPGSGTCEKCPDCACTEKECECKKFDRCALHDEMTNCRK-----RDEIES 641
 Db 623 CNCTDPEKFGPCEKCGTCLGVCARHKDCVOCRAFEK--BKETCSQCEMHFMNTRYES 680
 QY 642 VKEI-KDNGKDAV-NCFTKNEDDCVNRFQYEDSSGKILYVVEPECPKPDILVLLS 699
 Db 681 RKLTPQVHPDRPLSHCKEKVDGCFYFTYSNNGEASVHVEPECPSPDIIPIYAG 740
 QY 700 VNGAILLIGLALLIMKLLITHDRKEFAKEEERAKMDTANNPLYKEATSTFTNTY 759
 Db 741 VVAGIVLIGLALLIMKLLITHDRREPAKQSRERARERAMSNPLYKRPISHTYDTFNNKNSYNT 800
 QY 760 RG 761
 Db 801 EG 802

RESULT 13

fibronectin receptor beta chain precursor - human
 N:Alternate names: CD29 antigen; integrin beta-1 chain
 C:Species: Homo sapiens (man)
 C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 20-Aug-1999
 C:Accession: B27079
 R:Argaves, W.S.; Suzuki, S.; Arai, H.; Thompson, K.; Pierschbacher, M.D.; Ruoslahti, J. Cell Biol. 105, 1183-1190, 1987
 A:Title: Amino acid sequence of the human fibronectin receptor.
 A:Reference number: A27079; MUID:88007843
 A:Accession: B27079
 A:Molecule type: mRNA
 A:Residues: 1-798 <ARC>
 A:Cross-references: EMBL:X07979; NID:931441; PIDN:CAA30790.1; PID:931442
 C:Genetics:
 A:Gene: GDB:ITGB1; FNBR
 A:Cross-references: GDB:118732; OMIM:135630
 A:Map position: 10p11.2-10p11.2
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C:Keywords: alternative splicing; duplication; heterodimer; transmembrane protein
 F:1-20/Domains: signal sequence #status predicted <SIG>
 F:21-798/Product: fibronectin receptor beta chain #status predicted <MAM>

Query Match 44.1%; Score 1831; DB 2: Length 798;
 Best Local Similarity 44.9%; Pred. No. 9e-109;
 Matches 350; Conservative 134; Mismatches 268; Indels 28; Gaps 14;

QY 3 NICTTRGVSSCOCCIAVSPKAWCSD-----EALPLGSPRDLKENLKLKDNCAPESTFE 57
 Db 25 NKLCKANKSCGECIOAGPNCGWCTNSTFLOEGMPT-SARCDLDELAKKKGCPDIEDNP 83

F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-798/Product: integrin beta-1 chain #status predicted <MAT>

Query Match 43.8%; Score 1821; DB 2; Length 798;
Best Local Similarity 44.4%; Pred. No. 3, 9e-108;

Matches 346; Conservative 137; Mismatches 269; Indels 28; Gaps 13;

```
QY 3 NICTRGVSSCQCLAVSPMCANCS-----EALPLGSPRCDLKENILKDNCAPESTEPP 57
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 NRCLKANAKSCGECIAGPNCGWCTNTTFLQEGMT-SARCDLLEALKKKGCQPSDIENP 83
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58 VSEARVLEDRPLSDKSGDSQ-----VTQVSPRIALRLRPDSSKNESIQRQVEDYDV 112
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 84 RGSQTIKKKKNVNTNRKSKMAEKLRPEDITQIQOQLLKLRSGEPOKFTLKKFRADYPI 143
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 113 DIYYLMDLSYMKDLSIONLGTKLATQMRKLTSLRIGFAGFVDPKPSPYMYTISPPEA 172
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 144 DLYYLMDSLYSMKDLENVKSLGTDLMNEMRRTSDFRIGFSFVEKTYMPTISTTPAK- 202
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 173 LEMPCYDMKTTCLPAGFYKHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVC 232
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 203 LRNPC-TSQNCTSPESYNNVSLTDREGEFNEIVGOQRISGNLDSPEGGFDAIMQAVAC 261
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 233 DEKIGRNDSHLVETTPAKTHIALDGRLAGIVQPNQOQCHVGSNHSASTMDYPSL 292
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 262 GSLIGRN-VTRLVFPSTAGHFGADGKLGIVLPNDQCHL-ENNYTMSHYDYPSI 319
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 293 GLMTEKLSOKNINLFAVTENVNLYQNSSELPGTWGLSMDSSNVYQLIYDAVGKIR 352
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 320 AHLVOKLSENNIIQTLFAVTEEFQPYKELKNLIPKSAVGTLSGNSNVYQLIIDAVNSLS 379
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 353 SKVELEVRLPEELSLSPATCLN--NEYIPELKSCKMGIKIDIVSFSIEAKVRGCPQEK 410
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 380 SEVILNSKLPDGVITNKSYCKNGVNGTGENGRKCSNISIGDEVOFEISITANKCPNKE 439
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 411 EKSFYIKPVGFKDSLIVQVTFPDCACQAOAEPNSHRGNNGNGTPEGCVCRGCGPMLGSO 470
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 440 SETIKIKPLGTFEEVEYVLOFLCKNCOSHGIPASPKCHEGNGTFEGCGACRCNIEGRVGRH 499
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 CECSEEDYRPSQODE-CSPREGQPYVCSORGELCGOCVCHSSDFGK--ITGKYCECDDFS 527
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 500 CECSTDEVNSEMDAYCKRENSSEICSNNGECVCGQCVCRRKRDNTNIEYSGKFCEDNFN 559
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 528 CVRYGEMSGHGGSCGDDCLDSDMTGYCNCITTRDTCMSSNGLLCGRGKCEGSCV 587
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 560 CDRSGLICGGNGVCRVCECYPVYTGSCDCSLDTGPCLASNGOICNGRIGCEGACK 619
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 588 CIQPSYGDTCCKPTCPDACTFKKCECECKFDRGALHDENTCNRYCR---DEISVK-643
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 620 CTDPKFGPTCETCOTGLGVAEHKECVQCAFNKGEKD--TCAQECSHENLTKVESRE 677
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 644 ELKDTGK--DAVNCTYNEDCVYRFOYEDSSGKSLIYVEEPECPKGDILIVLLSYM 701
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 678 KLPQPVQDVPYTHCKEKDIDQCMFYFTYSVGNNEAIVHVEPDCPTGPIPIYAGV 737
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 702 GAILLIGLALLIKLITTIHDKREFAKFEERARAKMDTANPILYKEATSTFNITYRG 761
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 738 AGIYLIGLALLIKLIMITHDKREFAKFEERAKAKMDTGENDPIYKSAVTIVNPKYEG 797
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: May 19, 2002, 12:24:02
Job time: 9709 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 10:39:46 ; Search time 40.44 Seconds

(without alignments)
729,582 Million cell updates/sec

Title: US-09-673-302a-1

Perfect score: 4154
Sequence: 1 GPNICTTNGVSSCQCLAVS.....NNPLYKATSTFTNITYRGT 762

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4119	99.2	788	1	ITB3_HUMAN
2	3839	92.4	787	1	ITB3_MOUSE
3	2343	56.4	798	1	ITB3_MOUSE
4	2339	56.3	799	1	ITB5_HUMAN
5	2177	52.4	787	1	ITB6_MOUSE
6	2170.5	52.3	788	1	ITB6_HUMAN
7	1989.5	47.9	655	1	ITB5_HUMAN
8	1846.5	44.5	803	1	ITB5_MOUSE
9	1833	44.1	798	1	ITB1_CHICK
10	1831	44.1	798	1	ITB1_FELCA
11	1826.5	44.0	798	1	ITB0_HUMAN
12	1821	43.8	798	1	ITB1_MOUSE
13	1818.5	43.8	798	1	ITB1_XENLA
14	1817	43.7	773	1	ITB1_BOVIN
15	1801.5	43.4	799	1	ITB1_RAT
16	1676	40.3	577	1	ITB6_CAVPO
17	1579	38.0	809	1	PAT3_CAEEL
18	1537	37.0	845	1	ITB2_DROME
19	1520.5	36.6	769	1	ITB2_BOVIN
20	1515	36.5	771	1	ITB2_MOUSE
21	1513.5	36.4	769	1	ITB2_HUMAN
22	1509.5	36.3	769	1	ITB2_PIG
23	1506	36.3	806	1	ITB7_MOUSE
24	1502.5	36.2	798	1	ITB7_HUMAN
25	1331	32.0	769	1	ITB8_RABIT
26	1326.5	31.9	769	1	ITB8_HUMAN
27	1184	28.5	1822	1	ITB4_HUMAN
28	1154.5	26.8	1807	1	ITB4_RAT
29	1097	26.4	799	1	ITB1_DROME
30	260.5	6.3	4289	1	TENX_HUMAN
31	259	6.2	2201	1	TENX_HUMAN
32	255	6.1	1746	1	TENA_PIG
33	252.5	6.1	1808	1	TENA_CHICK

34	239.5	5.8	830	1	SREC_HUMAN
35	214	5.2	3106	1	LMH2_MOUSE
36	212	5.1	3672	1	LMH2_MOUSE
37	211	5.1	722	1	DL1_MOUSE
38	210	5.1	714	1	DL1_MOUSE
39	207	5.0	2437	1	NOTC_BRARE
40	205	4.9	2871	1	DL1_MOUSE
41	204	4.9	723	1	DL1_MOUSE
42	202	4.9	2531	1	NTCL_MOUSE
43	201.5	4.9	2139	1	NTCL_MOUSE
44	201	4.8	2871	1	FBM1_PIG
45	200.5	4.8	2444	1	NTCL_HUMAN

ALIGNMENTS

RESULT 1
ID ITB3_HUMAN STANDARD: PRT: 788 AA.
AC P05106: Q13413: Q16499: Q16448: Q15495;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Integrin beta-3 precursor (Platelet membrane glycoprotein IIb)
DE (GP1Ib) (CD61).
GN ITGB3 OR GP3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP MEDLINE-87165991; PubMed-3494014;
RA Fitzgerald L.A., Steiner B., Rall S.C. Jr., Lo S., Phillips D.R.;
RT "Protein sequence of endothelial glycoprotein IIb derived from a
RT cDNA clone, identity with platelet glycoprotein IIb and similarity
RT to 'Integrin'.";
RL J. Biol. Chem. 262:3936-3939(1987).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM BETA-3A).
RA MEDLINE-90265363; PubMed-2345548;
RA Fritchet P., Uzan G., Thevenon D., Denarter E., Prandini M.H.,
RA Marguerie G.;
RT "GPIIb and GPIIb amino acid sequences deduced from human
RT megakaryocyte cDNAs.";
RL Mol. Biol. Rep. 14:27-33(1990).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM BETA-3A).
RA MEDLINE-88213696; PubMed-2452834;
RA Zimrin A.B., Elisman R., Villalre G., Schwartz E., Bennett J.S.,
RA Poncez M.;
RT "Structure of platelet glycoprotein IIb. A common subunit for two
RT different membrane receptors.";
RL J. Clin. Invest. 81:1470-1475(1988).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM BETA-3A).
RA MEDLINE-90256778; PubMed-2341395;
RA Zimrin A.B., Gidwitz S., Lord S., Schwartz E., Bennett J.S.,
RA White G.C. II, Poncez M.;
RT "The genomic organization of platelet glycoprotein IIb.";
RL J. Biol. Chem. 265:8590-8595(1990).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM BETA-3C).
RA TISSUE-Osteoclastoma;
RC MEDLINE-97341177; PubMed-9195946;
RA Kumar C.S., James I.E., Wong A., Kwang V., Feld J.A.,
RA Nuthallaganti P., Connor J.R., Elchman C., Ali F., Hwang S.M.,
RA Riemann D.J., Drake F.H., Gowen M.;
RT "Cloning and characterization of a novel integrin beta3 subunit.";
RL J. Biol. Chem. 272:16390-16397(1997).
RN [6]
RP SEQUENCE OF 1-26 FROM N.A.

- RC TISSUE=Blood;
RX MEDLINE=94129007; PubMed=8298129;
RA Villa-Garcia M., Li L., Riely G., Bray P.F.;
RT "Isolation and characterization of a TATA-less promoter for the human
RT beta 3 integrin gene.";
RL Blood 83:668-676(1994).
- RN [7]
RX MEDLINE=91008291; PubMed=2145280;
RA Lanza F., Kiefer N., Phillips D.R., Fitzgerald L.A.;
RT "Characterization of the human platelet glycoprotein IIb gene.
RT Comparison with the fibronectin receptor beta-subunit gene.";
RL J. Biol. Chem. 265:18098-18103(1990).
- RN [8]
RX MEDLINE=93002753; PubMed=1382574;
RA Jlang W.M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,
RA Kristiansen G.W.;
RT "The gene organization of the human beta 7 subunit, the common beta
RT subunit of the leukocyte integrins HML-1 and LPM-1.";
RL Int. Immunol. 4:1031-1040(1992).
- RN [9]
RX MEDLINE=89313807; PubMed=2787511;
RA Van Kuppevelt T.H.M.S.M., Languino L.R., Gallit J.O., Suzuki S.,
RA Ruostelä E.;
RT "An alternative cytoplasmic domain of the integrin beta 3 subunit.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:5415-5418(1989).
- RN [10]
RX MEDLINE=87101510; PubMed=3801670;
RA Hiraiwa A., Matsukage A., Shiku H., Takahashi T., Naito K., Yamada K.;
RT "Purification and partial amino acid sequence of human platelet
RT membrane glycoproteins IIb and IIa.";
RL Blood 69:560-564(1987).
- RN [11]
RX MEDLINE=91158732; PubMed=2001522;
RA Calvete J.J., Henschen A., Gonzalez-Rodriguez J.;
RT "Assignment of disulphide bonds in human platelet GPIIb. A
RT disulphide pattern for the beta-subunits of the integrin family.";
RL Biochem. J. 274:63-71(1991).
- RN [12]
RX MEDLINE=96210016; PubMed=8631894;
RA Law D.A., Nannizzi-Alaimo L., Phillips D.R.;
RT "Outside-in integrin signal transduction. Alpha IIb beta 3-(GP IIb
RT IIb) tyrosine phosphorylation induced by platelet aggregation.";
RL J. Biol. Chem. 271:10811-10815(1996).
- RN [13]
RX MEDLINE=921482770; PubMed=11546839;
RA Xiong J.P., Stehle T., Dieffenbach B., Zhang R., Dunker R., Scott D.L.,
RA Joachimlak A., Goodman S.L., Arnout M.A.;
RT "Crystal structure of the extracellular segment of integrin alpha
RT Vbeta3.";
RL Science 294:339-345(2001).
- RN [14]
RX MEDLINE=89214713; PubMed=2565345;
RA Newman P.J., Derbes R.S., Aster R.H.;
RT "The human platelet alloantigens, PLAI and PLAI2, are associated with
RT glycoprotein IIb, and are distinguishable by DNA typing.";
RL J. Clin. Invest. 83:1778-1781(1989).
- RN [15]
RX MEDLINE=93055444; PubMed=1430225;
RA Wang R., Furlinata K., McFarland J.G., Friedman K., Aster R.H.,
RA Newman P.J.;
RT "An amino acid polymorphism within the RGD binding domain of platelet
RT membrane glycoprotein IIa is responsible for the formation of the
- RT Pena/Pemb alloantigen system.";
RL J. Clin. Invest. 90:2038-2043(1992).
- RN [16]
RX MEDLINE=93112977; PubMed=8093349;
RA Kuljpers R.W.A.M., Simsek S., Faber N.M., Goldschmeding R.,
RA van Weimerkerken R.K.V., von Dem Borne A.E.G.K.;
RT "Single point mutation in human glycoprotein IIa is associated with
RT a new platelet-specific alloantigen (Mo) involved in neonatal
RT alloimmune thrombocytopenia.";
RL Blood 81:70-76(1993).
- RN [17]
RX MEDLINE=94060373; PubMed=7694683;
RA Wang R., McFarland J.G., Kekomaki R., Newman P.J.;
RT "Amino acid 489 is encoded by a mutational 'hot spot' on the beta 3
RT integrin chain: the CA/TU human platelet alloantigen system.";
RL Blood 82:3386-3391(1993).
- RN [18]
RX MEDLINE=94179229; PubMed=8132570;
RA Santoso S., Kalb R., Kroll H., Walika M., Kiefel V.,
RA Mueller-Eckhardt C., Newman P.J.;
RT "A point mutation leads to an unpaired cysteine residue and a
RT molecular weight polymorphism of a functional platelet beta 3 integrin
RT subunit. The Sra alloantigen system of GPIIb.";
RL J. Biol. Chem. 269:8439-8444(1994).
- RN [19]
RX MEDLINE=90364410; PubMed=2392682;
RA Loftus J.C., O'Toole T.E., Plow E.F., Glass A., Frelinger A.L., III,
RA Ginsberg M.H.;
RT "A beta 3 integrin mutation abolishes ligand binding and alters
RT divalent cation-dependent conformation.";
RL Science 249:915-918(1990).
- RN [20]
RX MEDLINE=92156115; PubMed=1371279;
RA Bajt M.L., Ginsberg M.H., Frelinger A.L., III, Berndt M.C.,
RA Loftus J.C.;
RT "A spontaneous mutation of integrin alpha IIb beta 3 (platelet
RT glycoprotein IIb-IIa) helps define a ligand binding site.";
RL J. Biol. Chem. 267:3789-3794(1992).
- RN [21]
RX MEDLINE=92291320; PubMed=160206;
RA Lanza F., Stierle A., Fournier D., Morales M., Andre G., Nurdan A.T.,
RA Cazenave J.P.;
RT "A new variant of Glanzmann's thrombasthenia (Strasbourg I).
RT Platelets with functionally defective glycoprotein IIb-IIa complexes
RT and a glycoprotein IIa 214Arg-->214Tyr mutation.";
RL J. Clin. Invest. 89:1995-2004(1992).
- RN [22]
RX MEDLINE=93066201; PubMed=1438206;
RA Chen Y.-P., Djaffar I., Pliard D., Steiner B., Cleutaut A.-M.,
RA Caen J.P., Rosa J.-P.;
RT "Ser 752-->Pro mutation in the cytoplasmic domain of integrin beta 3
RT subunit and defective activation of platelet integrin alpha IIb beta
RT 3 (glycoprotein IIb-IIa) in a variant of Glanzmann thrombasthenia.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10169-10173(1992).
- RN [23]
RX MEDLINE=95184171; PubMed=8786622;
RA Bray P.F.;
RT "Inherited diseases of platelet glycoproteins: considerations for
RT rapid molecular characterization.";
RL Thromb. Haemost. 72:492-502(1994).
- RN [24]
RX MEDLINE=98025892; PubMed=9376589;
RA Basani R.B., Brown D.L., Vilaine G., Bennett J.S., Poncz M.;
RT "A Leu117-->Tyr mutation within the RGD-peptide cross-linking region


```

FT DISULFID 482 520 BY SIMILARITY.
FT DISULFID 487 496 BY SIMILARITY.
FT DISULFID 498 511 BY SIMILARITY.
FT DISULFID 526 531 BY SIMILARITY.
FT DISULFID 528 561 BY SIMILARITY.
FT DISULFID 533 546 BY SIMILARITY.
FT DISULFID 548 553 BY SIMILARITY.
FT DISULFID 567 572 BY SIMILARITY.
FT DISULFID 569 600 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
FT DISULFID 585 592 BY SIMILARITY.
FT DISULFID 606 611 BY SIMILARITY.
FT DISULFID 608 656 BY SIMILARITY.
FT DISULFID 613 623 BY SIMILARITY.
FT DISULFID 626 629 BY SIMILARITY.
FT DISULFID 633 642 BY SIMILARITY.
FT DISULFID 639 712 BY SIMILARITY.
FT DISULFID 660 688 BY SIMILARITY.
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 584 584 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MOD_RES 772 772 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 784 784 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 787 AA: 86694 MW: 86694 B150599ABC438A3 CRC64:

```

```

Query Match 92.4% Score 3839: DB 1: Length 787:
Best Local Similarity 91.3% Pred. No. 9, 7e-253:
Matches 694: Conservative 39: Mismatches 27: Indels 0: Gaps 0:

```

```

QY 3 NICTRGVSSCCQCLAVSPMCAMCSDALPLSPRCLEKLNKNCAPESIEFPVSEAR 62
DB 28 NICTRGVSSCCQCLAVSPMCAMCSDALPLSPRCLEKLNKNCAPESIEFPVSEAR 87
QY 63 VLEDRPLSKGSDSSQVTVQVFORALALRPDCKNSTIOYRVEDYVVDIYLMDSY 122
DB 88 ILEARPLSKGSDSSQVTVQVFORALALRPDCKNSTIOYRVEDYVVDIYLMDSY 147
QY 123 SMKDDMSIONLTGKATQMRKLTSLNRIGFGAFVDRKPSVWYISPPALNCPDMYT 182
DB 148 SMKDDMSIONLTGKATQMRKLTSLNRIGFGAFVDRKPSVWYISPPALNCPDMYT 207
QY 183 TCIPLMGFYHVLTLIDQVTRFNEEVKQSVSRNRDAPEGGFDAIMQATVDEKIGMRDA 242
DB 208 ACIPMGFYHVLTLIDQVTRFNEEVKQSVSRNRDAPEGGFDAIMQATVDEKIGMRDA 267
QY 243 SHLLVETPAKTHIALDGLAGIVQVNDGCHVSPNHNHSASTMDYPSLGMTLTKLSQK 302
DB 268 SHLLVETPAKTHIALDGLAGIVQVNDGCHVSPNHNHSASTMDYPSLGMTLTKLSQK 327
QY 303 NINLIFAVTENVNLQNTSELIPGTVGLVLSKSSNVQLIVDAVGRKRSVELEVRDL 362
DB 328 NINLIFAVTENVNLQNTSELIPGTVGLVLSKSSNVQLIVDAVGRKRSVELEVRDL 387
QY 363 PEBLSLSPNATCLINNEVIRGLKSCMGLKTGTVSFEISIAKRVGCPQEKESFTIRPVGRK 422
DB 388 PEBLSLSPNATCLINNEVIRGLKSCMGLKTGTVSFEISIAKRVGCPQEKESFTIRPVGRK 447
QY 423 DSLIYVTPDCCACQAOAQPNSHRNNGNTEFCGVCVCGGVMGWSQCECEEDYRPSQ 482
DB 448 DSLIYVTPDCCACQAOAQPNSHRNNGNTEFCGVCVCGGVMGWSQCECEEDYRPSQ 507
QY 483 QDECSREGOPVCSQREGELCGQCVCHSSDFKITGKYECODFSCVRRKSGMCGHGQC 542
DB 508 QDECSREGOPVCSQREGELCGQCVCHSSDFKITGKYECODFSCVRRKSGMCGHGQC 567
QY 543 SCGGDLCLSDMTGYCNCCTTRDTQMSNGLLCGRKCEGSCVCIOPGSGYGTCEKCP 602
DB 568 NCGDCVCDSDMTGYCNCCTTRDTQMSNGLLCGRKCEGSCVCIOPGSGYGTCEKCP 627
QY 603 TCPDACKTFRKEVECKKFPKRGALHDENTCNRCRDEISVSKELKOTGKDAVACTYKNEDD 662

```

```

DB 628 TCPDACKTFRKEVECKKFPKRGALHDENTCNRCRDEISVSKELKOTGKDAVACTYKNEDD 687
QY 663 CVRFRQYEDSSGKSLVVEEPECCKPDLIVLLSYGAILLIGLAALLIKMLLITTH 722
DB 688 CVRFRQYEDSSGKSLVVEEPECCKPDLIVLLSYGAILLIGLAALLIKMLLITTH 747
QY 723 DRKEFAKFEERAKKMDTANNPLYKATSTFTNTYKGT 762
DB 748 DRKEFAKFEERAKKMDTANNPLYKATSTFTNTYKGT 787

RESULT 3
ID ITB5_MOUSE STANDARD: PRT: 798 AA.
AC 070309; 070308; 088347.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Integrin beta-5 precursor.
GN ITGB5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS BETA-5A AND BETA-5B).
RC TISSUE=Liver.
RX MEDLINE=98198405; PubMed=9531507;
RA Zhang H., Tan S.M., Lu J.;
RT "CDNA Cloning reveals two mouse betas 5 integrin transcripts distinct in
RT cytoplasmic domains as a result of alternative splicing.";
RL Biochem. J. 331:631-637(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM BETA-5A).
RC TISSUE=Brain.
RX MEDLINE=98098874; PubMed=9880508;
RA Feng X., Tettebaum S.L., Quiroz M.E., Towler D.A., Ross F.P.;
RT "Cloning of the murine betas 5 integrin subunit promoter. Identification
RT of a novel sequence mediating granulocyte-macrophage colony-
RT stimulating factor-dependent repression of betas 5 integrin gene
RT transcription.";
RL J. Biol. Chem. 274:1366-1374(1999).
RN [3]
RP FUNCTION: INTEGRIN ALPHA-V/BETA-5 IS A RECEPTOR FOR FIBRONECTIN.
CC IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGAND.
CC -1 SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-5
CC ASSOCIATES WITH ALPHA-V.
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 ALTERNATIVE PRODUCTS: 2 ISOFORMS: BETA-5A (SHOWN HERE) AND BETA-
CC 5B. ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1 SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -1 SIMILARITY: CONTAINS 1 WFA-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF043257; AAC40110.1;
CC EMBL: AF043256; AAC40109.1;
CC EMBL: AF022110; AAD08782.1;
CC HSSP: P04355; 2MRT.
CC MGD: MGI:96614; Itgb5.
CC InterPro: IPR000561; BGF-like.
CC InterPro: IPR002369; Integrin_B.
CC InterPro: IPR003659; PSI.
CC InterPro: IPR002035; WFA.
CC Pfam: PF00362; Integrin_B_1.
CC PRINTS: PR01186; INTEGRINB.
CC ProDom: PD001811; Integrin_B_1.

```

DR SMART: SM00181; EGF; 1.
 DR SMART: SM00187; TNF; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 2.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
 DR Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Signal; Alternative splicing.
 FT SIGNAL 1 23
 FT CHAIN 1 23
 FT DOMAIN 24 798
 FT TRANSMEM 720 742
 FT DOMAIN 742 798
 FT DOMAIN 136 378
 FT DOMAIN 465 630
 FT REPEAT 465 512
 FT REPEAT 512 554
 FT REPEAT 554 593
 FT REPEAT 594 630
 FT DISULFID 463 463
 FT DISULFID 36 46
 FT DISULFID 39 75
 FT DISULFID 49 64
 FT DISULFID 202 211
 FT DISULFID 259 300
 FT DISULFID 401 413
 FT DISULFID 433 682
 FT DISULFID 461 465
 FT DISULFID 476 487
 FT DISULFID 484 522
 FT DISULFID 489 498
 FT DISULFID 500 513
 FT DISULFID 528 533
 FT DISULFID 530 563
 FT DISULFID 535 548
 FT DISULFID 550 555
 FT DISULFID 569 574
 FT DISULFID 571 602
 FT DISULFID 576 585
 FT DISULFID 587 594
 FT DISULFID 608 613
 FT DISULFID 610 657
 FT DISULFID 615 625
 FT DISULFID 628 631
 FT DISULFID 635 644
 FT DISULFID 641 714
 FT DISULFID 661 690
 FT CARBOHYD 347 347
 FT CARBOHYD 460 460
 FT CARBOHYD 479 479
 FT CARBOHYD 505 505
 FT CARBOHYD 586 586
 FT CARBOHYD 654 654
 FT CARBOHYD 705 705
 FT VARSPLIC 760 798
 FT
 FT CONFLICT 96 96
 FT CONFLICT 259 259
 FT CONFLICT 595 595
 FT SEQUENCE 798 AA: 87908 MW: 34898807E268880 CRC64:

Query Match 56.4%; Score 2343; DB 1; Length 798;
 Best Local Similarity 54.9%; Pred. No. 2, 4e-151;
 Matches 428; Conservative 129; Mismatches 199; Indels 24; Gaps 11;

QY 1 GPHICITRGVSSGQCCCLAVSPKACMSDEALPLGSP-----KCDLKENLTKNCAPESIE 55
 DB 24 GLNICTSGSATSCEDLLHPKACMSKEY--FGNPRSTSRCDLAKANLRNCGECE-IE 80
 QY 56 FVSEARVLEDRPLSDKG-SGDSQVTVQSPQIALRLRPDDSKNSIQVROVEDPVDI 114

DB 81 SPASSTHVLRLPLSSKSSATGSDVIOHTPOELVSLRPBEQTTFLQVQVQVEDYVDL 140
 QY 115 YYIMDLASYMKDMLSIONLGTIKLATQMRKLTSLNLIGFAGVDPKPYSPMYISPEALE 174
 DB 141 YYIMDLASLMKDDLENISLGTIKLAEMRKLTSLNRLGFGSFVDKDISPSFY-TAPRQT 199
 QY 175 NPC--YDMKTTCLPMFGYKHYVLTLDQVTRFNEEYKQKQSVSRNRDAPGCGDAIMQATVC 232
 DB 200 NCCIKYKFLPNKVPSPFGRHLLPLTRDVDSFNEEYKQKQSVSRNRDAPGCGDAVLOAVVC 259
 QY 223 DEKIGMRDASHLVFTTDAKHTALDGRLAGIYVQNGOCGHVSDNYSASTMDVPSL 292
 DB 260 KKKIKMRDASHLVFTTDDVPHIALDGLGGLVQPHDGHCHLMBANFTYASNOVDPSL 319
 QY 293 GLMTEKLSOKNINLFEATVENVNLQYVSELIPGTIVGVSIMDSNNLQILVDAYKIR 352
 DB 320 ALLGKLAENNNLFLFAVTKNHYMYKFTALIPOTTEILHDSKNLIQILINAYSSIR 379
 QY 353 SKVELLEVRDLPEELSLSFNATCLNNEVTPGLKSCMGKLGIDTVSFSLEAKYRGCP-QEKE 411
 DB 380 AKVELSVMDQPEDLNLFTATCODISYPCGRKCEGLKIGDTASFEYSVEARSCPGROAA 439
 QY 412 KSPETKPYGFKDLSLVQYTPDCCACQAOAPNSRCHNNGCTPECGYRCGPGMISQC 471
 DB 440 QSEFLRPVGRDLSLOVEAVYNCTGCGSTGLEPNSARQ-SGNGTYTGIGLCECDPGLGTRC 498
 QY 472 ECSEEDYRPSQDECSPPGCPVCSQREGLCGQCVCHSSPFCKITGKYCCDPSCCYRY 531
 DB 499 ECGEBENSGYNLCREAEGRKPLCGRGECSCNQCSCSESEFGRIYGFCCDPSFCARN 558
 QY 532 KGMCSGHQSCGQGLCDSDMTGYCNCRTTRTDQMSNGLLSCGRKCEGSCVCTOP 591
 DB 559 KGVLCSGHECHCGCKHAGYIGDNCNSFDVSTCAKAKDQICSDRGRVCVCGCQCTEP 618
 QY 592 GSYGPTGCKPCPACGFEKKECVCKKFRDGAHDETCNRKCRDELES-VKELKOTGK 650
 DB 619 GARGETCECPGCPACGSKKRCVCELLHOGK-PDNOTCHHOCDEXITVVDTVKDDQ 677
 QY 651 DAVNCTYRNEDDCVVRFOYEDDSGKSLIYVEEPECPKRPDILVLSVAGAILLGLA 710
 DB 678 EAVLCFYTAKCQVWVFSTELPNGRSLNLYLREPEGCSAPNANMTLLAVGSIILLGMA 737
 QY 711 ALLIKLLITTHDRKEPAKFEERAKAKMDTANPLLYKEATSTT-----NTTYGCT 762
 DB 738 LLAIKMLVLTTHDREPAKFQSESRARYEMASNPLYRKPISTHTVDFAFKFKNSYNGS 797
 RESULT 4
 ID ITB5_HUMAN STANDARD; PRT; 799 AA.
 AC P18084;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-5 precursor.
 GN ITB5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymic epithelium;
 RX MEDLINE=90228356; PubMed=2328726;
 RA Ramaswamy H., Hemler M.E.;
 RT "Cloning, primary structure and properties of a novel human integrin
 beta subunit";
 RL EMBO J. 9:1561-1568(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90319111; PubMed=2371275;
 RA Suzuki S., Huang Z.S., Tanthara H.;


```

Db 559 GSVGSGHGECHCECHAGYICDNCSTIDSTCRGRDQICSEHGCLCGCCTEP 618
Qy 592 GSYDTECEKPTCPDPACTFKKECECKKPRGALHDENTCNRCDIES-VKELDTGK 650
Db 619 GAFGECEKPTCPDPACTFKKECECKKPRGALHDENTCNRCDIES-VKELDTGK 650
Qy 651 DAVNCTKNEDECVRRQYVEDSGKSLVVEEPECPKDPDILVLLVSGAILLIGLA 710
Db 678 EAVLCFRTAKDCVMFTYVELPSGKSNLTVLREPECGNPNMTILLAVGSLILVGLA 737
Qy 711 ALTIWKLITITHDKEFAKFEERAKMDANPLKATSTFT-----NITVRC 762
Db 738 LLAIMKLVITIHDRERAKFOSEKRAKTEMASNPILKRPISHTVDFENKFNKSYNGT 797

RESULT 5
ID ITB6_MOUSE STANDARD; PRT: 787 AA.
AC 092079;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Integrin beta-6 precursor.
GN ITGB6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=20547423; PubMed=11095652;
RA Arend L.J., Smart A.M., Briggs J.P.;
RT "Mouse beta(6) integrin sequence, pattern of expression, and role in
RT kidney development.";
RL J. Am. Soc. Nephrol. 11:2297-2305(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-V/BETA-6 IS A RECEPTOR FOR FIBRONECTIN
CC AND CYTOTACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS
CC (BY SIMILARITY).
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-6
CC ASSOCIATES WITH ALPHA-V (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF115376; AAD17212.1; -
CC MGD: MGI:96615; Itgb6.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002369; Integrin_B.
DR InterPro: IPR001169; Integrin_Beta_C.
DR InterPro: IPR003659; PSI.
DR InterPro: IPR002035; VWFA.
DR Pfam: PF00362; Integrin_B_1.
DR PRINTS: PR01186; INTEGRINB.
DR ProDom: PD001811; Integrin_B_1.
DR SMART: SM00001; EGF_like_1.
DR SMART: SM00187; INB_1.
DR SMART: SM00423; PSI_1.
DR SMART: SM00327; VWFA_1.
DR PROSITE: PS00243; INTEGRIN_BETA_2.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
Repeat; Signal.

```

```

FT SIGNAL 1 21
FT CHAIN 22 787
FT DOMAIN 22 706
FT TRANSMEM 707 729
FT DOMAIN 730 787
FT DOMAIN 131 371
FT DOMAIN 456 619
FT REPEAT 502 543
FT REPEAT 544 582
FT REPEAT 583 619
FT DISULFID 23 454
FT DISULFID 31 41
FT DISULFID 34 70
FT DISULFID 44 59
FT DISULFID 197 204
FT DISULFID 252 293
FT DISULFID 384 406
FT DISULFID 426 469
FT DISULFID 452 456
FT DISULFID 467 479
FT DISULFID 476 511
FT DISULFID 481 490
FT DISULFID 492 502
FT DISULFID 517 522
FT DISULFID 519 552
FT DISULFID 524 537
FT DISULFID 539 544
FT DISULFID 558 563
FT DISULFID 560 591
FT DISULFID 565 574
FT DISULFID 576 583
FT DISULFID 597 602
FT DISULFID 599 645
FT DISULFID 604 614
FT DISULFID 617 620
FT DISULFID 624 633
FT DISULFID 630 701
FT DISULFID 649 677
FT CARBOHYD 48 48
FT CARBOHYD 97 97
FT CARBOHYD 260 260
FT CARBOHYD 387 387
FT CARBOHYD 418 418
FT CARBOHYD 463 463
FT CARBOHYD 471 471
FT CARBOHYD 541 541
FT CARBOHYD 575 575
SQ SEQUENCE 787 AA; 86041 MW; C6438C6F1E6B7FBD CRC64;

Query Match 52.4%; Score 2177; DB 1; Length 787;
Best Local Similarity 51.6%; Pred. No. 4,2e-140;
Matches 392; Conservative 124; Mismatches 232; InDels 12; Gaps 7;

Qy 5 CTTREYSSCOQCLAVSPMCACMSDE---ALPLGSRCLKENLKDNCAPESIEFPVSEA 61
Db 23 CAMGAEECSDDLGLGPPICACMSQENFTLGSAGERCPTPANLLAKGQLPFIENPVSR 82
Qy 62 RYLEDPLSDKSGDSQVOTGVSPORIALRLRPDQSKNFESQVQVEQYDPVDIYLDLS 121
Db 83 EYLQKPLPSVGRKNSDIOVAPOSILVKLRPGREGQLOVQVQTEQYDPVDIYLDLS 142
Qy 122 YSMKDLMSLONLGTKLATQMKLITSLNRIGFAGFVDPVSPYVYISPEALENCPYDMK 181
Db 143 ASMDLNTIKELGSRSLAKESKLTISNFRLGFSVPEKPSFPK-TTPEITWPCSSIP 201
Qy 182 TTCLEMPGKYKHYLTITDQVTRFNEVKKQSVSRNDAPEGGFALIMQATYVDEKIGRND 241
Db 202 YFCLETFGFKHLLPLTDAERFNEIVRKQISANIDPEGEFDLIMQAVCKEKIGRND 261
Qy 242 ASHLVFTTDAKTHALDGLAGIQQPNDGCHGSDHYSASTMDVPSLGLMTEKLSQ 301

```



```

FT DISULFID 538 573 BY SIMILARITY.
FT DISULFID 543 558 BY SIMILARITY.
FT DISULFID 560 565 BY SIMILARITY.
FT DISULFID 579 584 BY SIMILARITY.
FT DISULFID 581 612 BY SIMILARITY.
FT DISULFID 586 595 BY SIMILARITY.
FT DISULFID 597 604 BY SIMILARITY.
FT DISULFID 618 623 BY SIMILARITY.
FT DISULFID 620 666 BY SIMILARITY.
FT DISULFID 625 635 BY SIMILARITY.
FT DISULFID 638 641 BY SIMILARITY.
FT DISULFID 645 654 BY SIMILARITY.
FT DISULFID 651 728 BY SIMILARITY.
FT DISULFID 670 704 BY SIMILARITY.
FT MOD_RES 25 25 BLOCKED.
FT MOD_RES 788 788 PHOSPHORYLATION (BY TYR-KINASES)
(POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 525 525 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 589 589 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 803 AA; 88553 MW; 2F6FEFCDF2C80457 CRC64;

```

Query Match 44.5%; Score 1846.5; DB 1; Length 803;
 Best Local Similarity 45.1%; Pred. No. 1, 1e-117;
 Matches 353; Conservative 136; Mismatches 266; Indels 27; Gaps 15;

```

QY 1 GPNICTRGVSSCCOCLAVSPMCAMC--SDEALPLGSP---RCOLKENLMDNCAPESIEF 56
DQ 27 GGSPICIAMNKSCEGCEICQAGPNCWCKKTPFLDGEPTSRCDLALAKSCPEODIEN 86
QY 57 PVSARVLEDRPLSDKSGSSQ---VTVQSFORTIALRLRDPDSKNSIOVROVEDYP 111
DQ 87 PRGSRVLEDRREVYNRKIGAEKRLKPEALITQIOBKLVLRGVEPOTFSLKFRRAEDYP 146
QY 112 VDIYLLDLSYMKDDLMSTONIGTKALATPKRLTSRLRGFAPDKPSPVMTSPPE 171
DQ 147 IDLYLLMDLSYMKDDLMSTONIGTKALATPKRLTSRLRGFAPDKPSPVMTSPPE 171
QY 172 ALENPCYDMKTTCLPMFGYKHLVLTLDQVTRFNEEVYKQSVSRNRDPAPEGCFPAIMQATV 231
DQ 207 -LRNPGCG-DQNCSTSPSYKKNVLSLTSEGNKFNELVNGKQIISGNLDSPEGCFPAIMQAV 264
QY 232 CDEKIGRNDASHLLVFTTDAKTHIALDGLAGIIVPNDGQCHVSGDNHYSASTNDYPS 291
DQ 265 CGQDQIGNRN-VTRLLVFSTDAFGFHAGDGKLGIVLRPNDKCHL-ENNMVYTMHYDDPS 322
QY 292 LGLMTEKLSQKNINLIFAVENVVNLVXONSELIPGTIVVSMDSNNVQLIVDAKGI 351
DQ 323 IAHVQLSLNNIOTIFAVTEFOAVYKELKNLIPKSAVGLTSSNSNVQLIIDVAVSL 382
QY 352 RSKVLEVRDLPEELSLFNATCLN--NEVYPLKSCMGKIGDTYVSFSTEAKVRKCPQ 409
DQ 383 SSEVYIENSKLPKEVITISKYCKNGVNDIQEDGRKNSISIDEXRFEIANTANCPK 442
QY 410 -KKSFTIKRVGFKDLSIIVYTFDCCACQAOAEPRSHRCNNGNGTFEGGVCRGCGMUG 468
DQ 443 GQNETTIKIKLFTLEVEELHLOFCCLQSEEPSPACHPDNGFECCAGCNGRNGR 502
QY 469 SOCECESEEDYRPSQDE--CSPRGQPVCSORGECLCGCCCHSDPGK--ITGKVCQCD 525
DQ 503 RLECGSTDEVNSDMAYKARENSTEICSNGECIGQCVCKKRENTNEVYSKTECECN 562
QY 526 FSCVRYKGEKSGHGQSCGDCCLDSDMTGYCNCCTRTDTGCMSSNGLLSGRGKCEGS 585

```

```

DQ 563 FNCDRSNGLLCGGNGICKCYVCEFPNFTGSACDCLDTTPCWANGQJLCNGGTCECT 622
QY 586 CVCIOGSGVDDNCEKPTCPDCTAFKKEVECKKFPDRLAHDENTNRC-----RDEIS 641
DQ 623 CNETDPRFQGPTECMQOTCLGVCAEHRKDCVQCAPEK--EKEKETSQDCMHFNMRVES 680
QY 642 VKEL-KDTGKDAV-NCTYKNEDDCVVRFOYEDDSGKSLIYVEEPCPKPDILVLLS 699
DQ 681 RGLLPQVHDPDLSHCKENDVGCWFYFYVSNSEASVHVETECSPDILYVNG 740
QY 700 VMCALLIGLALLIKLLTTHDRKFAKFEERARANDVANNELVKEATSTFNITY 759
DQ 741 VVAGIVLIGLALLIKLLTTHDRKFAKFEERARANDVANNELVKEATSTFNITY 759
QY 760 RG 761
DQ 801 EG 802

RESULT 9
ID ITBL_FELCA STANDARD; PRT; 798 AA.
AC P53713:
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 01-MAR-2002 (rel. 41, Last annotation update)
DE Integrin beta-1 precursor (Fibronectin receptor beta subunit)
DE (CD29) (Integrin VLA-4 beta subunit).
GN ITGB1.
OS Fells silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Fells.
OX NCBI_TaxID:9685;
RN [1]
RP SEQUENCE FROM N.A.
RA Willett B.J.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-
CC 1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-
CC 1/BETA-1 AND ALPHA-2/BETA-1 RECOGNIZE THE PROLINE-HYDROXYLATED
CC SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-
CC 3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-6/BETA-1, ALPHA-
CC 10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR
CC FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN
CC THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN.
CC INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINOGEN. INTEGRIN
CC ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1
CC ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR
CC FOR VCAM1 AND RECOGNIZES THE SEQUENCE Q-T-D-S IN VCAM1. INTEGRIN
CC ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOFACTIN AND
CC OSTEOPOINTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-G-I-E-L IN
CC CYTOFACTIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPLIGRIN
CC AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR
CC VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A
CC WIDE ARRAY OF LIGANDS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1
CC ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-
CC 5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR
CC ALPHA-V.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U27351; AAC19407.1; -

```

DR InterPro: IPR000561: EGF-like.
 DR InterPro: IPR002369: Integrin_B.
 DR InterPro: IPR001169: Integrin_beta.C.
 DR InterPro: IPR003659: PSI.
 DR InterPro: IPR02035: VWFA.
 DR Pfam: PF00362: Integrin_B.1.
 DR PRINTS: PR01186: INTEGRIN_B.
 DR PRODOM: PD001811: Integrin_B.1.
 DR SMART: SM00001: EGF_Like.1.
 DR SMART: SM00187: INB.1.
 DR SMART: SM00423: PSI.1.
 DR SMART: SM00327: VWFA.1.
 DR PROSITE: PS00243: INTEGRIN_BETA.3.
 DR PROSITE: PS00022: EGF_1: UNKNOWN.2.
 KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Signal; Phosphorylation.
 FT CHAIN 1 20
 FT DOMAIN 21 798
 FT TRANSMEM 21 728
 FT DOMAIN 729 751
 FT DOMAIN 752 798
 FT DOMAIN 140 378
 FT DOMAIN 466 635
 FT REPEAT 466 515
 FT REPEAT 516 559
 FT REPEAT 560 598
 FT REPEAT 599 635
 FT DISULFID 27 464
 FT DISULFID 35 45
 FT DISULFID 38 75
 FT DISULFID 48 64
 FT DISULFID 207 213
 FT DISULFID 261 301
 FT DISULFID 401 415
 FT DISULFID 435 691
 FT DISULFID 462 466
 FT DISULFID 477 489
 FT DISULFID 486 525
 FT DISULFID 491 500
 FT DISULFID 502 516
 FT DISULFID 531 536
 FT DISULFID 533 568
 FT DISULFID 538 553
 FT DISULFID 555 560
 FT DISULFID 574 579
 FT DISULFID 576 607
 FT DISULFID 581 590
 FT DISULFID 592 599
 FT DISULFID 613 618
 FT DISULFID 615 661
 FT DISULFID 620 630
 FT DISULFID 633 636
 FT DISULFID 640 649
 FT DISULFID 646 723
 FT DISULFID 665 699
 FT MOD_RES 783 783
 FT CARBOHYD 50 50
 FT CARBOHYD 94 94
 FT CARBOHYD 97 97
 FT CARBOHYD 212 212
 FT CARBOHYD 269 269
 FT CARBOHYD 363 363
 FT CARBOHYD 406 406
 FT CARBOHYD 417 417
 FT CARBOHYD 481 481
 FT CARBOHYD 520 520
 FT CARBOHYD 584 584
 FT CARBOHYD 669 669
 SQ SEQUENCE 798 AA: 88092 MW: 241C3871046D838 CRC64:

Query Match 44.1%; Score 1833; DB 1; Length 798;
 Best Local Similarity 44.6%; Pred. No. 8.7e-11;

Matches 348; Conservative 136; Mismatches 268; Indels 28; Gaps 13;
 QY 3 NICTTGAVSSCQCLAVSPWCAMCSD-----EALPLGSPRCULKEKLKDNCAPSIEEP 57
 DB 25 NRIKLANAKSGGICGAGPNCWCNVSTFLQSBEMPL-SARCDLLEALKKGGCHPDIEEP 83
 QY 58 VSEARVLEDRPLSDKSGDSQ-----VTVQSPORIALRLRPDSKNFSIOVRQVEDPY 112
 DB 84 RGSKDYKKNKNVNNRSGKTGTEKLPEDINQIQOQVLVLDRSGEPQFTLKFRAEDYPI 143
 QY 113 DIYIMDLSTSMKDLMSIONLCTKATLQMRKLTSLNRIGFAPVUKPSPMYISPPA 172
 DB 144 DLYIMDLSTSMKDLMSIONLCTKATLQMRKLTSLNRIGFAPVUKPSPMYISPPA 202
 QY 173 LEPDCTYMKTTCLPMGYKHLVLTDOVTFENVEVKOSYSRRRDAPFEGFADIMQATVC 232
 DB 203 LRNFC-TSENCSTSPSYKAVLSLTDKGEVFNLVGKRISGULDSDFEGFADIMQAVNC 261
 QY 233 DEKIGWRNDASHLVFTTDAKTHIALDGLAGIVQPNQGHVSDNNHYSASTMDYPSL 292
 DB 262 GSLIGWEN-VTRLVFTSDAGFHFAGDGKLGIVLPNDGCHLEND-VYTMSHYDYPSI 319
 QY 293 GLMTERLSQKNINLIPAVTENYVNLQNSBELIPGTGVGLMSDSSNVQLIYDAYGXR 352
 DB 320 AHIWQKLSENNIOPIFAVTEEPQYKELKNLIPKSAVGLSANSSNVQLIYDAYNISLS 379
 QY 353 SKVELEVRLPEELSLSFNATCLN--NEVYIPGLSKMGKIGDTVSFSEAKVRCPOBK 410
 DB 380 SEVLENSKLPPEGVITISYKCKNGVNGENRCKNSISGDEVQFELSTANKCPNNK 439
 QY 411 EKSEITIPVGRKSLIVQVTFDDCACQQAEPNSHRCKNNKCTFECCVGRCPGLGSG 470
 DB 440 SEIIRKIPLFTFEVEIILQIFICECQNGIIPSSPCHEGNSFCGACRKEGVNGH 499
 QY 471 CEGSEEDYRPSQODE--OSPREGQPVCSRGECGCGCVCCHSPDFGR--ITGKCECDPS 527
 DB 500 CECSTDEVSEMDDAYTKENSSIFGSCNNEGCVCQGVCKRNTNMTISGKCECDN 559
 QY 528 CVRYKMGSGHGGSCGDCGLCDSDMTGYCNCCTTRTDCMSSNGILSGRGKCEGSCV 587
 DB 560 CDRSNGILCGANGVCRCVCECNPNYGSACDSLDPTSCMATNGOICNRRGICEGACK 619
 QY 588 CIOPGSYGDCCEKCPDCTCFKCEVECKKDRGALHDEMCNRCR---DEIESYK 643
 DB 620 CTDRKFGPCCEMOTLGVCAEHKEVCQCRAFKGGKD--TCAQFCSHFNITKENRD 677
 QY 644 ELKDTGR--DAVNCTYNEEDCVRFQYEDSSGSLIYVEEPCEPKGDIIVLLSVN 701
 DB 678 KLPPGQVDPLSHCKEKDQVDCWFYFTYSVNGNEALVHVVEPCEPTGPDIIPIVAGVY 737
 QY 702 GAILLIGLALLWKLLITTHDKREKAFEEERARAKMDTANNPLKFASTFTNTIYRG 761
 DB 738 AGIYVILGIALLLIWKLLIMTHIDREFAKFEKKNAMKMDGENPIYKSAVTVVNPKEG 797
 RESULT 10
 ID ITBI_HUMAN STANDARD; PRT; 798 AA.
 AC POS556; P78466; P78467; Q13089; Q14647; Q13090; Q13212; Q13091;
 AC Q14623;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-1 precursor (Fibronectin receptor beta subunit)
 DE (CD29) (Integrin VLA-4 beta subunit).
 GN ITGB1 OR FNBR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN 111
 RP SEQUENCE FROM N.A. (ISOFORM BETA-1A).
 RC TISSUE=Placenta;

RX MEDLINE=8807843; PubMed=2958481;
 RA Argaves W.S., Suzuki S., Arai H., Thompson K., Pletschbacher M.D.,
 RA Rueschli E.;
 RT "Amino acid sequence of the human fibronectin receptor.";
 RL J. Cell Biol. 105:1183-1190(1987).
 RN [2]
 RP SEQUENCE OF 717-757 FROM N.A. AND CHARACTERIZATION OF BETA-1B.
 RX MEDLINE=93209984; PubMed=7681433;
 RA Balzac F., Belkin A.M., Koteliarsky V.E., Balabanov Y.V., Altruda F.,
 RA Silengo L., Tarone G.;
 RT "Expression and functional analysis of a cytoplasmic domain variant of
 the beta 1 integrin subunit.";
 RL J. Cell Biol. 121:171-176(1993).
 RN [3]
 RP SEQUENCE OF 717-757 FROM N.A. AND FUNCTION.
 RX MEDLINE=95014744; PubMed=7523423;
 RA Balzac F., Rette S.F., Albini A., Melchiorri A., Koteliarsky V.E.,
 RA Geuna M., Silengo L., Tarone G.;
 RT "Expression of beta 1b integrin isoform in CHO cells results in a
 dominant negative effect on cell adhesion and motility.";
 RL J. Cell Biol. 127:557-565(1994).
 RN [4]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS BETA-1A; BETA-1B; BETA-1C AND
 BETA-1D).
 RX TISSUE-SKELETAL MUSCLE;
 MEDLINE=95398646; PubMed=7545396;
 RA Zhidkova N.I., Belkin A.M., Mayne R.;
 RT "Novel isoform of beta 1 integrin expressed in skeletal and cardiac
 muscle.";
 RL Biochem. Biophys. Res. Commun. 214:279-285(1995).
 RN [5]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1B).
 RX TISSUE-SKELETAL MUSCLE;
 MEDLINE=91065539; PubMed=2249781;
 RA Altruda F., Cervella P., Tarone G., Balzac F., Stefanuto G.,
 RA Silengo L.;
 RT "A human integrin beta 1 subunit with a unique cytoplasmic domain
 generated by alternative mRNA processing.";
 RL Gene 95:261-266(1990).
 RN [6]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1D).
 RX MEDLINE=95377431; PubMed=7544298;
 RA van der Flier A., Kulkman I., Baudoin C., van der Neut R.,
 RA Sonnenberg A.;
 RT "A novel beta 1 integrin isoform produced by alternative splicing:
 unique expression in cardiac and skeletal muscle.";
 RL FEBS Lett. 369:340-344(1995).
 RN [7]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS BETA-1A AND BETA-1C).
 RX TISSUE-CERVICAL CARCINOMA;
 MEDLINE=92202279; PubMed=1551917;
 RA Langino L.R., Rueschli E.;
 RT "An alternative form of the integrin beta 1 subunit with a variant
 cytoplasmic domain.";
 RL J. Biol. Chem. 267:7116-7120(1992).
 RN [8]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1C-2).
 RX MEDLINE=98161805; PubMed=9494054;
 RA Slieng G., Pressler R., Johansson S.;
 RT "Identification of beta1c-2, a novel variant of the integrin beta1
 subunit generated by utilization of an alternative splice acceptor
 site in exon C.";
 RL Biochem. J. 330:1255-1263(1998).
 CC -1- FUNCTION: INTEGRIN ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-
 1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-
 1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED
 SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-
 3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-8/BETA-1, ALPHA-
 10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR
 FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN
 THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN.
 INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRONECTIN. INTEGRIN
 ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1

CC ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR
 CC FOR VCAM1. IT RECOGNIZES THE SEQUENCE O-I-D-S IN VCAM1. INTEGRIN
 CC ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOTACTIN AND
 CC OSTROPOINTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-G-I-E-L IN
 CC CYTOTACTIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPLIGRIN
 CC AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR
 CC VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A
 CC WIDE ARRAY OF LIGANDS. ISOFORM BETA-1B INTERFERES WITH ISOFORM
 CC BETA-1A RESULTING IN A DOMINANT NEGATIVE EFFECT ON CELL ADHESION
 CC AND MIGRATION (IN VITRO).
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1
 CC ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-
 CC 5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR
 CC ALPHA-V.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ISOFORM BETA-1B
 CC DOES NOT LOCALIZE TO FOCAL ADHESIONS.
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: BETA-1A (SHOWN HERE), BETA-1B,
 CC BETA-1C, BETA-1C-2 AND BETA-1D; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: ISOFORM BETA-1A IS WIDELY EXPRESSED. OTHER
 CC ISOFORMS ARE GENERALLY COEXPRESSED WITH A MORE RESTRICTED
 CC DISTRIBUTION. ISOFORM BETA-1B IS EXPRESSED IN SKIN, LIVER,
 CC SKELETAL MUSCLE, CARDIAC MUSCLE, PLACENTA, UMBILICAL VEIN
 CC ENDOTHELIAL CELLS, NEUROBLASTOMA CELLS, LYMPHOMA CELLS, HEPATOMA
 CC CELLS AND ASTROCYTOMA CELLS. ISOFORMS BETA-1C AND BETA-1C-2 ARE
 CC EXPRESSED IN MUSCLE, KIDNEY, LIVER, PLACENTA, CERVICAL EPITHELIUM,
 CC UMBILICAL VEIN ENDOTHELIAL CELLS, FIBROBLAST CELLS, EMBRYONAL
 CC KIDNEY CELLS, PLATELETS AND SEVERAL BLOOD CELL LINES. ISOFORM
 CC BETA-1C-2, RATHER THAN ISOFORM BETA-1C, IS SELECTIVELY EXPRESSED IN
 CC PRIMARY T-CELLS. ISOFORM BETA-1C IS EXPRESSED IN NONPROLIFERATING
 CC AND DIFFERENTIATED PROSTATE GLAND EPITHELIAL CELLS. ISOFORM BETA-
 CC 1D IS EXPRESSED SPECIFICALLY IN STRIATED MUSCLE (SKELETAL AND
 CC CARDIAC MUSCLE).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTRAINS 1 WMPA-LIKE DOMAIN.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD29 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd29.htm".
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; X07979; CAA30790.1; -;
 CC DR EMBL; U33882; AAA79835.1; -;
 CC DR EMBL; U33879; AAA79835.1; JOINED.
 CC DR EMBL; M84237; AAA74402.1; -;
 CC DR EMBL; U33879; AAA79832.1; -;
 CC DR EMBL; M34189; AAA59182.1; -;
 CC DR EMBL; U33880; AAA79833.1; -;
 CC DR EMBL; U33879; AAA79833.1; JOINED.
 CC DR EMBL; U28252; AAA81366.1; -;
 CC DR EMBL; U33882; AAA79834.1; -;
 CC DR EMBL; U33879; AAA79834.1; JOINED.
 CC DR EMBL; U33881; AAA79834.1; JOINED.
 CC DR EMBL; M84237; AAA74403.1; -;
 CC DR PIR; B27079; B27079.
 CC DR HSSP; P15358; 1SKZ.
 CC DR MIM; 135630; -;
 CC DR InterPro; IPR000561; EGF-like.
 CC DR InterPro; IPR002369; integrin_B.
 CC DR InterPro; IPR001169; integrin_beta.C.
 CC DR InterPro; IPR003659; PSI.
 CC DR InterPro; IPR002035; WFA.
 CC Pfam; PF00362; integrin_B.1.
 CC PRINTS; PR01186; INTEGRIN.
 CC PRODOM; PD001811; integrin_B.1.
 CC SMART; SM00001; EGF-like; 1.
 CC SMART; SM00187; INB; 1.
 CC SMART; SM00423; PSI; 1.

DR SMART: SM00327; VWA: 1.
 DR PROSITE: PS00243; INTEGRIN_BETA: 3.
 DR PROSITE: PS00242; EGF_1; UNKNOWN: 2.
 DR Integrin; Cell adhesion; Receptor; Transmembrane; glycoprotein;
 KW Repeat; Signal; Phosphorylation; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 798 INTEGRIN BETA-1.
 FT DOMAIN 21 728 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 729 751 POTENTIAL.
 FT DOMAIN 752 798 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 140 378 VFMA-LIKE.
 FT DOMAIN 466 635 4 CYSTEINE-RICH TANDEM REPEATS.
 FT REPEAT 466 515 I.
 FT REPEAT 516 559 II.
 FT REPEAT 560 598 III.
 FT REPEAT 599 635 IV.
 FT DISULFID 27 464 BY SIMILARITY.
 FT DISULFID 35 45 BY SIMILARITY.
 FT DISULFID 38 75 BY SIMILARITY.
 FT DISULFID 48 64 BY SIMILARITY.
 FT DISULFID 207 213 BY SIMILARITY.
 FT DISULFID 261 301 BY SIMILARITY.
 FT DISULFID 401 415 BY SIMILARITY.
 FT DISULFID 435 466 BY SIMILARITY.
 FT DISULFID 462 466 BY SIMILARITY.
 FT DISULFID 477 489 BY SIMILARITY.
 FT DISULFID 486 525 BY SIMILARITY.
 FT DISULFID 491 500 BY SIMILARITY.
 FT DISULFID 502 516 BY SIMILARITY.
 FT DISULFID 531 536 BY SIMILARITY.
 FT DISULFID 533 568 BY SIMILARITY.
 FT DISULFID 538 553 BY SIMILARITY.
 FT DISULFID 555 560 BY SIMILARITY.
 FT DISULFID 574 579 BY SIMILARITY.
 FT DISULFID 576 607 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 FT DISULFID 592 599 BY SIMILARITY.
 FT DISULFID 613 618 BY SIMILARITY.

Query Match 44.1%; Score 1831; DB 1; Length 798;
 Best Local Similarity 44.9%; Pred. No. 1.2e-116;
 Matches 350; Conservative 134; Mismatches 268; Indels 28; Gaps 14;

QY 3 NICTRGVSSCCQCLAVSPMCACSD-----EALPLGSPRCIDKKNLKNDCAPESIEFP 57
 DB 25 NNCGLAANKSCGCECIQAGPCNCGMCTNSFTLQEGMT-SARCDDEALKKKCGPPDIDNP 83
 QY 58 VSEARVLEDRPLSDKSGDSSQ-----VYQVSPRIALRLRPDSKNFSIOVROVEDYPV 112
 DB 84 RGSKDIIKKKNVTNRSKGTAEKLPEDIHQIOPQOLVLRSGEPQTFLEKRAEDYPI 143
 QY 113 DIYYLMDLSYSKMDLMSLQNLGTLATQMKLTSLNLRIGPAFADKRVVSPMYISTPPEA 172
 DB 144 DIYYLMDLSYSKMDLMSLQNLGTLATQMKLTSLNLRIGPAFADKRVVSPMYISTPPEA 202
 QY 173 LKNPCYDMKTTCLPMEGYKHVLTLPDOVTRFNEEKKOSVSHNRDAPGEGDAIMQATVC 232
 DB 203 LKNPC-TSQNCTTPPSYKNTSLNKGVEFELVQKORISNLDSPEGEPALMQAVCA 261
 QY 233 DERIGRNDASHLLVFTTDAKTHIALDGLAGIYVOPNDGQCHVSDNHYSASTMDYVSL 292
 DB 262 GSLIGWRN-VTRLVFTSTDAFGHAGDGKIGIVLPNDQCHL-NNMYTMSHYDYPSI 319
 QY 293 GLMTELSKKNINLFAVTEENVNLYKQNSLIPCTYGVVLSMDSNNVQLIYDAVYKIR 352
 DB 320 AHLVOKLSNNIQTIFAVTEERQPYVKELKNLIPASVAGTSANSSNYIQLIIDYNSLS 379
 QY 353 SKVELEVRDLPEELSLSFNATCLN-NEVYPLKSCMGKIDYVFSFIEAKVRCPQEK 410
 DB 380 SEVILENGKLSREVITISYCKNGVNGTEGMRKCSNLSIDVQFELSTISNKKCPKKD 439
 QY 411 EKSFITKPVGFKDSLIVQVTFDQDCAQQAQAPNSHRCNNGNGTEGVCRCGPGMLDSQ 470

DB 440 SDFKIRPLIGFTEVEVILLQYICECEQSEGIPEBSPKHEGNGTEGACRNEGRVGRH 499
 QY 471 CESEEDYRPSQDE-CSPREQPVCSQREGLCGQCVCHSSDFGK--ITGKYCCDDPS 527
 DB 500 CECSDEVNSEDADACRKNSEISQNSNNECEGCGVCVRKRDNTNREISGFCEDNFN 559
 QY 528 CVRRKEMCSGHHGQSCGDLCDSDMTGYCNCCTPTDTCMSSNGLCSGRKCEGSCV 587
 DB 560 CDPSNGLICGNVCKRCRCVCECPNNTGSACDCSISDTSTCEASNGICNGRICEGVCK 619
 QY 588 CIQPGSYGDTCEKCPDPCACTPKKCEVCKEFGKALHDEMTNRYCR----DELESVK 643
 DB 620 CTDPKFGQGTCEMCQTCGLGVCAEHKECVQCARFNKEKRD--TCTQECYSFNTYESKD 677
 QY 644 EL-KDQGXDAV-NCTYKNEDDCVRFQYEDSSGSIIVVEEPECKPGDILVLSVM 701
 DB 678 KLPQVQPPDVSHSKREDVDDCFYTTYSNGNNEVAVHVEPECPIDPIIPIVAGV 737
 QY 702 GAILLIGLALLIMKLITITDRKEPAKFEERARAKMDTANNPLYKATSFNTIYRG 761
 DB 738 AGVILGLALLIMKLITITDRKEPAKFEERARAKMDTANNPLYKATSFNTIYRG 797

RESULT 11
 ITBO_XENLA STANDARD; PRT; 798 AA.
 ID ITBO_XENLA STANDARD; PRT; 798 AA.
 AC P12607;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-1* precursor.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP MEDLINE=8816829; PubMed=2833505;
 RA Desimone D.W., Hynes R.O.;
 RT "Xenopus laevis integrins. Structural conservation and evolutionary
 divergence of integrin beta subunits.";
 RL J. Biol. Chem. 263:5333-5340(1988).
 CC -1- FUNCTION: BETA INTEGRINS ASSOCIATE WITH ALPHA SUBUNITS TO FORM
 RECEPTOR COMPLEXES THAT RECOGNIZE THE SEQUENCE R-G-D IN A WIDE
 CC ARRAY OF LIGANDS (BY SIMILARITY).
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY. THE
 CC INTEGRINS BETA-1 AND BETA-1* OF XENOPUS ONLY DIFFER BY 19 AMINO
 CC ACIDS.
 CC -1- SIMILARITY: CONTAINS 1 VFMA-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M20180; AAA4890.1; -.
 DR PIR: B28193; B28193.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001168; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR002035; VFMA.
 DR Pfam: PF00362; Integrin_B_1.
 DR Pfam: PF01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B_1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.

Db	144	LYYLMDLSFSKDDLENNKSLGTLAMTMEKMTSDPFRIGFSGFVEKTVMPPISTTPAKLL	203
Qy	174	ENPCYDMKTTCTLPMEGKXHVLTLDQYTRNEBVKKOSVSNRPDARECGEDATQAFVCD	233
Db	204	-NPCIN-DONCTSPFYSKYVNLTKDKGLFNLDLGKOOISGNLSDSPGCGFALIQVAPVC	261
Qy	234	EKIGRNDASHLVPTPDAKTHILDLGLAITVOPNGOOCVGSDDNYSASTWIDYSLG	293
Db	262	EOIGRNR-VTLLVLFSTDAGHFRAGDGLGSLVINDGRCLH-GNNMTYSHYDYDSIA	319
Qy	294	LMTEKLSQKNLNLFFAVYENVVNLVQNSSELPCTPVGVLSMLBSSNYQLDLVADYKIRS	353
Db	320	HLVQKLSNNIQTIFAVYEDPQPYQDELKNLIPRSANVTGSSNSNVYQLDLIDYSLSS	379
Qy	354	KVELEVRDLPEBELSLSEFNATCLNNEVITGL- -KSCGNLKIULTVYSLEAKVRCSPQE	409
Db	380	ELLINENSKLPEGVITISYRSFCKNG- -VKGTEGDEGRKCSNISJQGVPELISVTAHKCPKK	437
Qy	410	-KENSFTTKPYGFDLSLVQYTPDDCCACQADRPNSHRCKNKGNTTECCVCRGPGWLG	466
Db	438	GOAESIKIKPIGFMEEVYIVLQFTCECDQDQGTNPSECFHFGTTECGACRCNBSRIG	497
Qy	469	SOCSESEEDYPPSOODE-CSPREGOPVCSRGECILCGOCVCHSSDFGK- -TTGKYCECD	522
Db	498	KECESCTDEVASDEMDAYCRENSENSEISNNGDCLCGOCYCKKADNPENYSGKTYCECDN	557
Qy	526	FSCVARKGEMCSGHHOCSGCDLSDMTGYGCNCTTPTDCKNSNGLSGRNCCEGS	585
Db	558	FNCDRSNGLIGCGKGYCVRVCECPNYSGSACDSEDTSCMAKNQICNGRIGCOGR	611
Qy	586	CVCIDPGSYGDTCEKCPRPCACTFFKACVCECKKFFDGRALHDEYTCNRYCR- -DELES	644
Db	618	CKCDPDKRGQGTCELQCVGVTCEHKKCYOCRAFOKEKOD- -VCMEQCHMFIISLVDS	675
Qy	642	VKELKDTG- -KDAVNCTYKKNEDCVPRROYEYEDSSGKSIILVVEEPRCKPRDILVLLS	695
Db	676	REELPQPGQAEALTHCKRKADMECHFTYTVSDSKENVMVAVVVEPRPSPDILPIYAG	733
Qy	700	VMGATLILGLAALLIWLKILITIHOKREKAFKEEERAPAKMDTANPLYKEATSTFTNTY	755
Db	736	VVAGVILIGLALLILWLKILITIHDRREPAKFEKEKEMAKMDTGEMPIYKSAVATVPNKY	795
Qy	760	RG 761	
Db	796	EG 797	
RESULT 12			
ITBL_MOUSE	ITBL_MOUSE	STANDARD;	PRT; 798 AA.
AC	P09055;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-NOV-1988 (Rel. 09, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Integrin beta-1 precursor (fibrinectin receptor beta subunit) (CD29) (Integrin VLA-4 beta subunit).		
GN	ITGB1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRIN-BALB/C;		
RC	MEDLINE=89055707; PubMed=3262537;		
RA	Tomihara S.;		
RT	"Murine mRNA for the beta-subunit of integrin is increased in		
RT	BALB/c-3T3 cells entering the G1 phase after the G0 state."		
RL	FEBS Lett. 238:315-319(1988).		
RN	[2]		
RP	SEQUENCE OF 2-798 FROM N.A.		
RC	STRIN-BALB/C;		
RC	MEDLINE=89235580; PubMed=2523953;		

Query Match	Best Local Similarity	Score	DB 1:	Length
Matches 346; Conservative 137; Mismatches 269; Indels 28; Gaps 13.	44.4%;	1821.	DB 1:	798;
3 NICTRGVSSCGOCLAVSPMCANCSQ-----FALDLSGPRDLKENTLKNCAPESEIEPP 57	43.8%;	1821.	DB 1:	798;
25 NRCLKANAKSGGCGICDAGPWCAGCTVTTFLDGGMT-SARCDLLEALKKGCQSPDIENE 83	44.4%;	1821.	DB 1:	798;
58 VSEARVLEDPRLSDKSGDSQ-----VYQSPQRIALRLPDDSKNFESIQVROVEDYVP 112	44.4%;	1821.	DB 1:	798;
84 RGSOTIKKNNVYNRKSGMAEKLREPDIQIQPOGLLLKLSGPRQKFTLKFRAEDYPI 143	44.4%;	1821.	DB 1:	798;
113 DIYIMDLSTSMKDDLSIQNLGTYKLAFTQMRKLSNLRIGFAGFVDRPVSPYNISSPEA 172	44.4%;	1821.	DB 1:	798;
144 DLYIMDLSTSMKDDLENYKSLGTDLMMEMRRTISDFRIGSGSEVEKTVMPYISTTPAK- 202	44.4%;	1821.	DB 1:	798;
173 LENCYIMAKTTCPLMGYGYKHVLTLDQVYTRNEPEYKQKQSVSRNDADEGGFDALIMATYC 232	44.4%;	1821.	DB 1:	798;
203 LRNPG-TSEQNCISPSFYKNVLSLTDREFNELVLGGQRIISGNLDSPEGGFDALIMOVAVC 261	44.4%;	1821.	DB 1:	798;
233 DEKIGMENDSHLYVTTDAKTHIALDRLAGIYQPNDCQCHVSDNHYSASTTMDYPSL 292	44.4%;	1821.	DB 1:	798;
262 GSLIGMRN-VTRLIVSTDAEGHFGADGKLGIVLPNDGCHL-ENNVTYMSHTYDPST 319	44.4%;	1821.	DB 1:	798;
293 GLMTEKLSQKNINLIPAVTENVNLYQNSSELIDPTQIVGLSDMSSNVLLIDVAGKIR 352	44.4%;	1821.	DB 1:	798;


```

FT DISULFID 412 668 BY SIMILARITY.
FT DISULFID 443 443 BY SIMILARITY.
FT DISULFID 454 446 BY SIMILARITY.
FT DISULFID 463 502 BY SIMILARITY.
FT DISULFID 468 477 BY SIMILARITY.
FT DISULFID 478 493 BY SIMILARITY.
FT DISULFID 508 513 BY SIMILARITY.
FT DISULFID 510 545 BY SIMILARITY.
FT DISULFID 515 530 BY SIMILARITY.
FT DISULFID 532 537 BY SIMILARITY.
FT DISULFID 551 556 BY SIMILARITY.
FT DISULFID 553 584 BY SIMILARITY.
FT DISULFID 558 367 BY SIMILARITY.
FT DISULFID 569 576 BY SIMILARITY.
FT DISULFID 590 595 BY SIMILARITY.
FT DISULFID 592 638 BY SIMILARITY.
FT DISULFID 597 607 BY SIMILARITY.
FT DISULFID 610 613 BY SIMILARITY.
FT DISULFID 617 626 BY SIMILARITY.
FT DISULFID 623 700 BY SIMILARITY.
FT DISULFID 642 676 BY SIMILARITY.
FT MOD.RES 760 760 PHOSPHORYLATION (BY SIMILARITY).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 773 773
SQ SEQUENCE 773 AA; 85322 MM; 54340886CE157195 CRC64;

```

Query Match 43.7%; Score 1817; DB 1; Length 773;
 Best Local Similarity 44.2%; Pred. No. 1e-115;
 Matches 344; Conservative 138; Mismatches 268; Indels 28; Gaps 13;

```

QY 3 NICTTREVSSCOCCLAVSPMCAMCSD-----EALPLGSPDCDLKENLKNCAPESEEP 57
DB 2 NRCIKANAKSCGECIQAGPCMGCTNSTFLQEGMPT-SARCDLEALKKKGCHPDIENP 60
QY 58 VSEARVEDRPLSPKSGSSD-----VQVSPQRIALRLRPDSKSFISQVROVEDYPV 112
DB 61 RGSFOIKKNNKVTNRSGTAEKLPQEDITQIQPOLVLRSGEPQFTFLKRAEDYPT 120
QY 113 DLYVLMDSYMKDDLSIONLGTKLATQMRKLTSLNLRIGGAFVDPKPVYVYISPPPA 172
DB 121 DLYVLMDSYMKDDLSIONLGTKLATQMRKLTSLNLRIGGAFVDPKPVYVYISPPPA 172
QY 173 LNPFCYMKTTCLPMEGYKHYVLTDOVYTRPNEVKKQSSRNDAPEGFDALIMQATVC 232
DB 180 LRNPCCTN-EQNCSTSPFSYKYNVLSLTDGEVFNELVGRKQRTSGNLDSPGEGFDALIMQAVC 238
QY 233 DEKGMNDASHLLVETTTDAKTHALRAGIYQPNDDGCHGSDNHVASTMIDPSTL 292
DB 239 GSLIGMNN-VTRLLVETSTDAFHAGGKGLGIVLPDGGCHLEND-VYMSHYIYDPT 296
QY 293 GLMTEKLSOKNINLIFAVTENNVNLYONYSELAPGTTGVGLSMDSNVTLQIVDAYGKIR 352
DB 297 AHLVOKLSENNIQTIFAVTEEFQVYVYKELKLPKSAVGLTANSSNVLIIDAYNSLS 356
QY 353 SKVELEPRDLPBELSLFNATCLN-NEVLPGLKSCMGKLTGIVVSIIEAKYRGCPQEK 410
DB 357 SEVILENSKLABEGVTINYSKYCKNGVNGTEGNGKCSNISIGDVOFEISITANKCPKN 416
QY 411 EKSTFIPVGFKSLIYQVFPDCACAOAEPNHRGNGNGNGNGFEGVCGRGGWLGSQ 470
DB 417 SETIKIRPLGTEVEVILQPTICBCEGCEITPOSPPKCHGNGTIFEGCACRNGRGRH 476

```

```

QY 471 CECSEEDYRPSQODE-CSPREGQVYCSORGELCGQCVCHSSDFGK--LTGKYCECDDES 527
DB 477 CECSTDEVNSEMDYAYCHKENSSEICSNNGECVQCQCRDRDNTNELYSKPECECNFN 536
QY 528 CVRYKGEKSGHGGCSCGDDCLDSDMTGYVGNCTRTDTCSSNGLLCSGKCEGSCV 587
DB 537 CDRSNGLLGGGNGVCKVNCBPNYTGACDCSLGTTSCMAVNGQICNGVCECGACK 596
QY 588 CIOPESSYDTECEKPTCPDCTPFKKEVECKKFPDGLAHENTGNRYCR---DEISVK 643
DB 597 CNDPFGQPTCEMOCOTGVCAGHEKCEYQCAFPNKEKD--TCAQECSHLNTIKVENRD 654
QY 644 ELKDTGK--DAVNCTYKNEDDCVVRFQYEDSSGKSLIYVVEEPCEPGLIVLVSVM 701
DB 655 KLPQGGQVPLSHCKEKEDVDCEMFYFTSYVNGNNAATVAVHETCEPGLIPIVAGV 714
QY 702 GALLIGLALLIWLTLTTHDRKEFAFEERAKADTANNPYKATSTFNITY 759
DB 715 AGVILGLALLIWLTLTTHDRKEFAFEERAKADTANNPYKATSTFNITY 772

```

```

RESULT 15
ID ITBI_RAT STANDARD: PRT: 799 AA.
AC P49134.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Integrin beta-1 precursor (Fibronectin receptor beta subunit)
DE (CD29) (Integrin VLA-4 beta subunit).
GN ITGB1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9531632; Pubmed=7541764;
RA Malek-Hedayat S., Rome L.H.;
RT "Cloning and sequence of the cDNA encoding the rat oligodendrocyte
  integrin beta 1 subunit.";
RL Gene 158:287-290(1995)
CC -1- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-
  1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-
  1/BETA-1 AND ALPHA-11/BETA-1 RECOGNIZE THE PROLINE-HYDROXYLATED
  SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-
  3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-8/BETA-1, ALPHA-
  10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR
  FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN
  THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN.
  INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINOGEN. INTEGRIN
  ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1
  ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR
  FOR VCAM1 AND RECOGNIZES THE SEQUENCE C-I-D-S IN VCAM1. INTEGRIN
  ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOTOXIN AND
  OSTEOPOINTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-C-I-E-L IN
  CYTOTOXIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPLIGRIN
  AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR
  VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A
  WIDE ARRAY OF LIGANDS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1
  ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-
  5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR
  ALPHA-V.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WMA-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isd-sib.ch/announce/>
CC or send an email to license@isd-sib.ch).

DR	EMBL: U12309; AAA86669.1; -	
DR	InterPro: IPR000561; EGF-like.	
DR	InterPro: IPR002369; Integrin_B.	
DR	InterPro: IPR001169; Integrin_beta_C.	
DR	InterPro: IPR003659; PSI.	
DR	InterPro: IPR002035; VFMA.	
DR	Pfam: PF00362; Integrin_B; 1.	
DR	PRINTS: PR01186; INTEGRINB.	
DR	ProDom: PD00181; EGF_like; 1.	
DR	SMART: SM00001; EGF_like; 1.	
DR	SMART: SM00187; INB; 1.	
DR	SMART: SM00423; PSI; 1.	
DR	SMART: SM00327; VMA; 1.	
DR	PROSITE: PS00243; INTEGRIN_BETA; 3.	
DR	PROSITE: PS00022; EGF_1; UNKNOWN; 2.	
KM	Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;	
KW	Repeat; Signal; Phosphorylation.	
FT	SIGNAL	1
FT	CHAIN	21
FT	DOMAIN	21
FT	TRANSMEM	730
FT	DOMAIN	753
FT	DOMAIN	799
FT	DOMAIN	140
FT	DOMAIN	467
FT	REPEAT	517
FT	REPEAT	517
FT	REPEAT	561
FT	REPEAT	599
FT	REPEAT	600
FT	DISULFID	27
FT	DISULFID	35
FT	DISULFID	38
FT	DISULFID	48
FT	DISULFID	207
FT	DISULFID	261
FT	DISULFID	401
FT	DISULFID	435
FT	DISULFID	463
FT	DISULFID	478
FT	DISULFID	487
FT	DISULFID	492
FT	DISULFID	501
FT	DISULFID	532
FT	DISULFID	533
FT	DISULFID	534
FT	DISULFID	539
FT	DISULFID	556
FT	DISULFID	575
FT	DISULFID	577
FT	DISULFID	582
FT	DISULFID	593
FT	DISULFID	614
FT	DISULFID	616
FT	DISULFID	621
FT	DISULFID	631
FT	DISULFID	634
FT	DISULFID	637
FT	DISULFID	641
FT	DISULFID	647
FT	DISULFID	666
FT	MOD_RES	784
FT	CARBOHYD	50
FT	CARBOHYD	94
FT	CARBOHYD	97
FT	CARBOHYD	212
FT	CARBOHYD	269
FT	CARBOHYD	363
FT	CARBOHYD	406
FT	CARBOHYD	417
FT	CARBOHYD	482
FT	CARBOHYD	521
FT	CARBOHYD	585
FT	CARBOHYD	670
DR	EMBL: U12309; AAA86669.1; -	
DR	InterPro: IPR000561; EGF-like.	
DR	InterPro: IPR002369; Integrin_B.	
DR	InterPro: IPR001169; Integrin_beta_C.	
DR	InterPro: IPR003659; PSI.	
DR	InterPro: IPR002035; VFMA.	
DR	Pfam: PF00362; Integrin_B; 1.	
DR	PRINTS: PR01186; INTEGRINB.	
DR	ProDom: PD00181; EGF_like; 1.	
DR	SMART: SM00001; EGF_like; 1.	
DR	SMART: SM00187; INB; 1.	
DR	SMART: SM00423; PSI; 1.	
DR	SMART: SM00327; VMA; 1.	
DR	PROSITE: PS00243; INTEGRIN_BETA; 3.	
DR	PROSITE: PS00022; EGF_1; UNKNOWN; 2.	
KM	Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;	
KW	Repeat; Signal; Phosphorylation.	
FT	SIGNAL	1
FT	CHAIN	21
FT	DOMAIN	21
FT	TRANSMEM	730
FT	DOMAIN	753
FT	DOMAIN	799
FT	DOMAIN	140
FT	DOMAIN	467
FT	REPEAT	517
FT	REPEAT	517
FT	REPEAT	561
FT	REPEAT	599
FT	REPEAT	600
FT	DISULFID	27
FT	DISULFID	35
FT	DISULFID	38
FT	DISULFID	48
FT	DISULFID	207
FT	DISULFID	261
FT	DISULFID	401
FT	DISULFID	435
FT	DISULFID	463
FT	DISULFID	478
FT	DISULFID	487
FT	DISULFID	492
FT	DISULFID	501
FT	DISULFID	532
FT	DISULFID	533
FT	DISULFID	534
FT	DISULFID	539
FT	DISULFID	556
FT	DISULFID	575
FT	DISULFID	577
FT	DISULFID	582
FT	DISULFID	593
FT	DISULFID	614
FT	DISULFID	616
FT	DISULFID	621
FT	DISULFID	631
FT	DISULFID	634
FT	DISULFID	637
FT	DISULFID	641
FT	DISULFID	647
FT	DISULFID	666
FT	MOD_RES	784
FT	CARBOHYD	50
FT	CARBOHYD	94
FT	CARBOHYD	97
FT	CARBOHYD	212
FT	CARBOHYD	

SQ SEQUENCE 799 AA; 88494 MW; F4475202EB8A3FA6 CRC64;

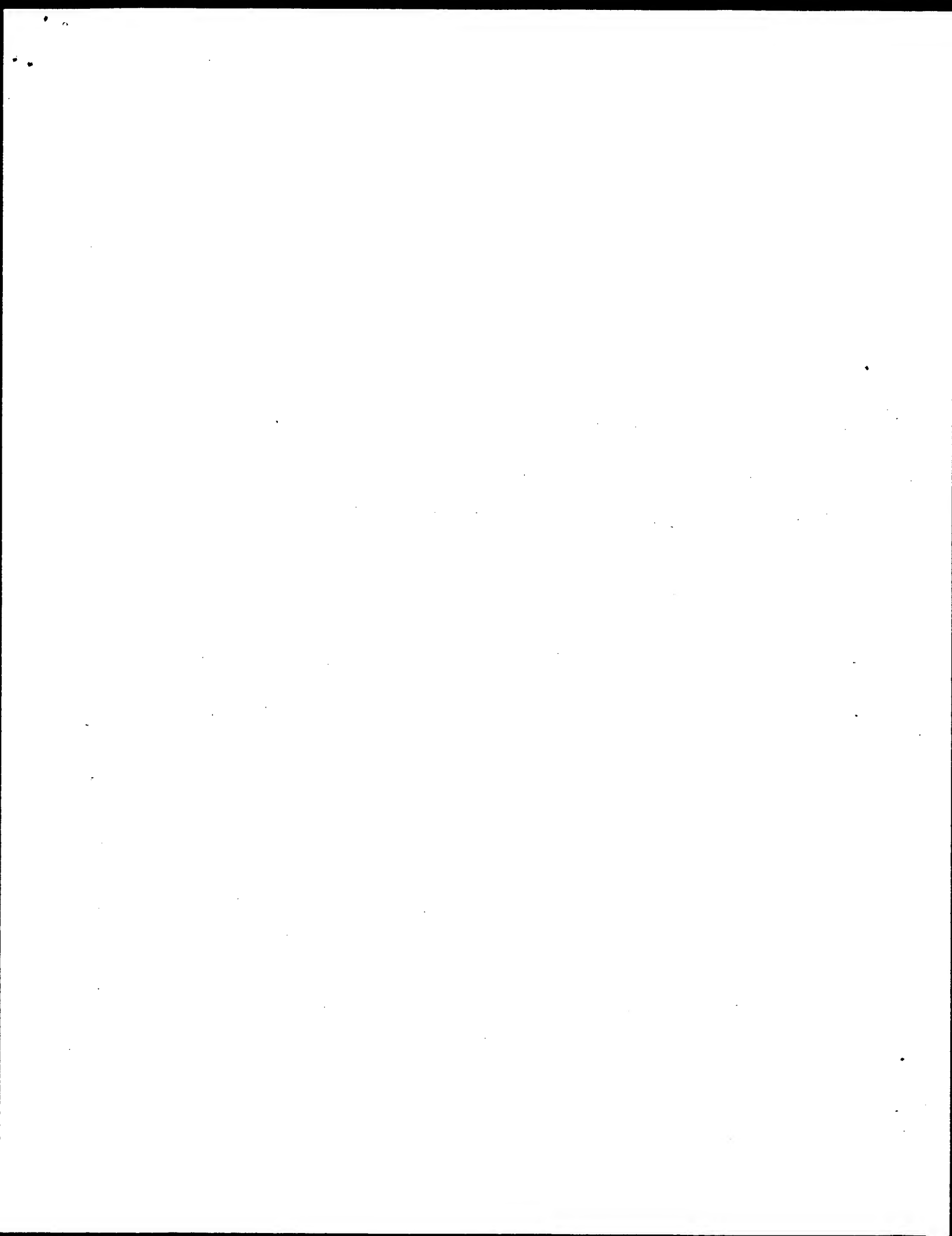
Query Match

Best Local Similarity 44.28; Pred. No. 1.2e-114.

Matches 345; Conservative 136; Mismatches 271; Indels 29; Gaps 14;

QY	3	NICTGTVSGCOOCLAVSPKAWCSD-----EALPGSPKCOLKEULLKDNCAPEISIEFP	57
Db	25	NRLCLANAKASGCEBCTIOAGBNGCWCTINTTFLIOEGMPT-SACDDLEMLKKGKCHPSDIEBP	83
QY	58	VSEARVLEDRPLSDKSGDSSQ-----VTOVSPORALRLRPDSKNFSIYQVQVEDPV	114
Db	84	RGSQITKKNNKNTNSRKGMAEKLPEDITQIOPQOLLRLSRGEPKFTLFRKREADYPI	145
QY	113	DIYVLMDSYSKKDDLMSTIONGTRKLTQWRKLTSLNRIGFAFYDKPVSPTMTISPEA	174
Db	144	DLVYLMDSYSKKDDLENNKSLGTDLNEMKRITSDIRIGFSGFEVKYIMPISTTPAK-	202
QY	173	LENDPCYDMKTTCCLPMFGYKHWLTLDDQVTRNEEVVKOSYSRNRDAPREGFPAIMQAVC	233
Db	203	LBNPC-TSDQNTSPFSYKKNVLSLSDGEFFNELYGOQRSLSGULDSRGGPRAIMQAVC	264
QY	233	DEKIGKNDASHLIVFTTDAKTHALDGRAGIYQPNDSQCHVSDNHYASATIMDYPSTL	292
Db	262	GSLIGRRN-VTRLEFSTVDAGFHEACGKLGIGIYLPNDQCHL-ENNYWTMSHYDAPSI	319
QY	293	GLMTEKLSOKNINLFAATEWVNLXONSELIPTGTIVGLMSDSNNVLOLIVAYGKIR	352
Db	320	AHLVQLSNNINQITLFAATEEPQVYELKNLIPKAVGLSIGSSNNVLOLIDAYNSLS	379
QY	353	SKVELEVRDLPEBLSYSTNATCLN-NEVJIPGLKSCMGKIGIPVFSFEAKVGC-PQE	409
Db	380	SEVITLNSKLPDGVITNKYSYCKNGVGTGBENRKCSNISIDEVOFELSTIAKCPNKE	439
QY	410	KESFTIKRVGKSDLIYQVTFDDCAQOAEPNHRCONNGNITFEBCVGRCPGMLGS	468
Db	440	SEMOULKNLNGLEVEVEVLOFICKKCSOHHGIPAPKCHENGTFEGACRCEGRVGR	499
QY	470	QCSCSEEDYRPSQODE-CPSPREGPVCYSQGBECLQCCYCHSDSK--ITGKVCCEDDF	526
Db	500	HGCSSTDEVNSDMAYORAKRNSSEISNNGBEQCCQYCKRRENTNITYSGKCECDNF	559
QY	527	SCVIRKGEKSGHGQSCGDCCLSDSDMTGYCNCITTRDTCKSSNGLLCSRGKCEGSC	566
Db	560	NCDRSNGLICGNGVCRWCECYPNTGSGACDCLDTPVCATNGQJCNRGICCEGCAC	619
QY	587	VCLQPSYGDPCCEPCPCPDACITFKKCBVCBCKKFDGALHDMTQCNRCR----DELESV	642
Db	620	KCIDRPFQGPTECQTLQGVCAEKKHCYOCRAFNENGERKD--TQADCSHFNLTKYESR	677
QY	643	KEIKDPTGK-DAVNCQTYKNEDCCVVRPOYEDSSGSLIYVEEPCFKGDIJVLVLSV	700
Db	678	EKLQPPQVQDPTVTHCKEKDDCFYFTTYSVNSKGAHNVETLPTDCTGDIIPYAGV	737
QY	701	MGAIIILGIALAILMKLITLIDRKKEFAKFEERARAKWTDANPDLXEATSPITNTYR	760
Db	738	VAGIYVLIGALLILMKLITLIDHREBAKFEKERNMAKMDGENPTRYSAVTVVNPKE	797
QY	761	G 761	
Db	798	G 798	

Search completed: May 19, 2002, 12:27:36
Job time: 6470 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 10:38:58 ; Search time 96.49 Seconds
(without alignments)
1366.174 Million cell updates/sec

Title: US-09-673-302A-1
4154
Sequence: 1 GPNICTRGVSSCQCLAVS.....NPLYEKATSTFNITVRCGT 762

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organella:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-ivirna:*
16: SP-bacteriap:*
17: SP-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4016	96.7	784	6	097702	097702 Canis famill
2	4013	96.6	784	6	097707	097707 Canis famill
3	3996	96.2	788	6	097703	097703 Canis famill
4	3898	93.8	784	6	0957H1	0957H1 sus scrofa
5	3894	93.7	784	6	097705	097705 sus scrofa
6	3606	86.8	723	11	096W16	096W16 ratius sp.
7	3484.5	83.9	781	13	092071	092071 gallus gall
8	3472	83.6	680	11	096W15	096W15 mus sp. bet
9	3260.5	75.5	788	13	007012	007012 xenopus lae
10	2315.5	55.7	791	6	096K49	096K49 bos taurus
11	1833	44.1	798	6	096GP0	096GP0 sus scrofa
12	1800.5	43.3	807	13	097A01	097A01 ictalurus p
13	1658	39.9	788	5	096444	096444 blomphalari
14	1609	38.7	837	5	09NAs7	09NAs7 anophelies g
15	1588.5	38.2	783	5	P92163	P92163 strongyloce
16	1584	38.1	799	5	Q95P95	Q95P95 crassostrea

17	1553	37.4	772	13	092070	092070 gallus gall
18	1508	36.3	806	5	061677	061677 lytechinus
19	1506	36.3	772	13	09PUD4	09PUD4 ictalurus p
20	1505	36.2	806	5	P91808	P91808 strongyloce
21	1498	36.1	764	6	097704	097704 canis famill
22	1498	36.1	802	5	076727	076727 strongyloce
23	1420.5	34.2	767	13	0987H6	0987H6 cyprinus ca
24	1403.5	33.8	768	13	0987H8	0987H8 cyprinus ca
25	1390.5	33.3	792	5	017494	017494 acropora ml
26	1382	33.5	801	5	P91774	P91774 pacifastacu
27	1312	31.6	790	5	09GSE3	09GSE3 podocoryne
28	1225.5	29.5	640	11	009182	009182 ratius norv
29	1145	27.6	878	5	097189	097189 geodia cydo
30	1080.5	26.0	738	11	088424	088424 mus musculu
31	1079.5	26.0	838	5	018482	018482 ophittaspon
32	1064	25.6	764	5	097343	097343 suberites d
33	828.5	19.9	439	4	09BUC9	09BUC9 homo sapien
34	754.5	18.2	290	11	09D214	09D214 mus musculu
35	738	17.8	150	11	063846	063846 ratius sp.
36	603	14.5	134	6	096M51	096M51 sus scrofa
37	419	10.1	194	13	091415	091415 gallus gall
38	417.5	10.1	403	4	014549	014549 homo sapien
39	417.5	10.1	494	4	095965	095965 homo sapien
40	368	8.9	69	6	0957C4	0957C4 bos taurus
41	347	8.4	103	11	P97483	P97483 mus musculu
42	327.5	7.9	166	6	018961	018961 bos taurus
43	318	7.7	65	4	012806	012806 homo sapien
44	304	7.3	114	11	09R151	09R151 cavia porce
45	281	6.8	94	6	Q95N85	Q95N85 ovls arles

ALIGNMENTS

RESULT 1
ID 097702 PRELIMINARY: PRT: 784 AA.
AC 097702:
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PLATELET GLYCOPROTEIN IIA.
GN GPIIb.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99410043; PubMed=10482317;
RA Lipcomb D.L., Bourne C., Boudreau M.K.;
RT "DNA sequence of the canine platelet beta3 gene from cDNA: comparison
RT of canine and mouse beta3 to segments that encode alloantigenic sites
RT and functional domains of beta3 in human beings.";
RL J. Lab. Clin. Med. 134:313-321(1999).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC EMBL: AF116270; AAD13680.1; .
CC InterPro: IPR002086; Aldehyde_dehydr.
CC InterPro: IPR000561; EGF-like
CC InterPro: IPR002369; Integrin_B.
CC InterPro: IPR001169; Integrin_B.
CC InterPro: IPR003659; PSI.
CC InterPro: IPR002035; VWFA.
CC Pfam: PF00362; Integrin_B.1.
CC PRINTS: PR01186; INTEGRINB.
CC PRODOM: PR001811; Integrin_B.1.
CC SMART: SM00001; EGF_Like; I.
CC SMART: SM00147; INB; I.
CC SMART: SM00423; PSI; I.
CC SMART: SM00327; VWA; I.

DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 3.
 KM Cell adhesion: Cytoskeleton; EGF-like domain; Extracellular matrix;
 KM Glycoprotein; Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 784 AA; 86388 MW; 6B2A6D34916EA260 CRC64;

Query Match 96.7%; Score 4016; DB 6; Length 784;
 Best Local Similarity 96.3%; Pred. No. 0;
 Matches 732; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 3 NICTRGVSSCQCCCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEFPVSEAR 62
 DB 25 NICTRGVHSCQCCCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCALESIIEFPVSEAR 84
 QY 63 VLEDRPLSDKSGSDSSQVTVQVSPQIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDLSY 122
 DB 85 ILLEAPPLSNKSGSDSSQITQVSPQIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDLSY 144
 QY 123 SMKDDLSIONLGTKLATQMRKLTSLNLRIGGAFVDPKPSYMYISPEALENFCYDMKT 182
 DB 145 SMDDLSIONLGTKLASQMHKLTSLNLRIGGAFVDPKPSYMYISPEALENFCYDMKT 204
 QY 183 TCLPMFGYKHVLTLDQVTRFNEEVKQSVSRNRDAPEGGFDAIMQATVDEKIGMRNDA 242
 DB 205 TCLPMFGYKHVLTLDQVTRFNEEVKQSVSRNRDAPEGGFDAIMQATVDEKIGMRNDA 264
 QY 243 SHLVFTTDAKTHIALDGRLAGIVQPDGQCHVSDNHSASTTMDYPSLGLMTEKLSQK 302
 DB 265 SHLVFTTDAKTHIALDGRLAGIVQPDGQCHVSDNHSASTTMDYPSLGLMTEKLSQK 324
 QY 303 NINLIFAVTEENVNLXONYSELPFTTVGVLSMSSNVQLIYDAVGKIRSKVELEVRDL 362
 DB 325 NINLIFAVTEENVNLXONYSELPFTTVGVLSMSSNVQLIYDAVGKIRSKVELEVRDL 384
 QY 363 PEEILSLFNATCLNNEVTPGLKSCMGKLTIGDIVSFSIEAKVRGCPQEKESFTIKPVGFR 422
 DB 385 PEEILSLFNATCLNNEVTPGLKSCMGKLTIGDIVSFSIEAKVRGCPQEKESFTIKPVGFR 444
 QY 423 DSLIYQVTFDCCACQAOAEPNHRNNGNGTEFCGVCAGGWMLSQCECSEEDYRPSQ 482
 DB 445 DSLIYQVTFDCCACQAOAEPNHRNNGNGTEFCGVCAGGWMLSQCECSEEDYRPSQ 504
 QY 483 QDECSPREGQPVCSQREGLCGQCVCHSSDPFGKITGKYCEDDFSCVARKGEMSGHGQC 542
 DB 505 QDECSPREGQPVCSQREGLCGQCVCHSSDPFGKITGKYCEDDFSCVARKGEMSGHGQC 564
 QY 543 SCGDCLCDSMDWGYCNCCTTRDTFCMSSNGLLCSGRKCEGSCVCIOPGSYGDTCEKCP 602
 DB 565 SCGDCLCDSMDWGYCNCCTTRDTFCMSSNGLLCSGRKCEGSCVCIOPGSYGDTCEKCP 624
 QY 603 TCGDACTFKKECECKECPRGALHDENTNCRKCRDIESEYKELDTGKRAVNCITYKNEDD 662
 DB 625 TCGDACTFKKECECKECPRGALHDENTNCRKCRDIESEYKELDTGKRAVNCITYKNEDD 684
 QY 663 CVARFQYEDSSGKSLIYVEEPECKGPDILVLLSVGAILLIGLALLIMKLLITTH 722
 DB 685 CVARFQYEDSSGKSLIYVEEPECKGPDILVLLSVGAILLIGLALLIMKLLITTH 744
 QY 723 DRKEFAKFEERARAKMDTANNPLKYKATSTFTNITYRGT 762
 DB 745 DRKEFAKFEERARAKMDTANNPLKYKATSTFTNITYRGT 784

RESULT 2
 Q9TUN7 PRELIMINARY; PRT; 784 AA.
 AC Q9TUN7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 13, Last annotation update)

DE GLYCOPROTEIN GPIIIA.
 GN GPIIIA.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA Tao J., Parrilla R.;
 RT "Molecular cloning of dog GPIIIA cDNA."
 RL Submitted (Jul-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL, AF170525; AADA9737.1;
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001659; Integrin_beta_C.
 DR InterPro: IPR002035; vWFA.
 DR Pfam, PF00362; Integrin_B; 1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom, PD001811; Integrin_B; 1.
 DR SMART: SM00001; EGF-like; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PST; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 3.
 KM Cell adhesion: Cytoskeleton; EGF-like domain; Extracellular matrix;
 KM Glycoprotein; Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 784 AA; 86416 MW; 9D507F827628790A CRC64;

Query Match 96.6%; Score 4013; DB 6; Length 784;
 Best Local Similarity 96.2%; Pred. No. 0;
 Matches 731; Conservative 15; Mismatches 14; Indels 0; Gaps 0;

QY 3 NICTRGVSSCQCCCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEFPVSEAR 62
 DB 25 NICTRGVHSCQCCCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCALESIIEFPVSEAR 84
 QY 63 VLEDRPLSDKSGSDSSQVTVQVSPQIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDLSY 122
 DB 85 ILLEAPPLSNKSGSDSSQITQVSPQIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDLSY 144
 QY 123 SMKDDLSIONLGTKLATQMRKLTSLNLRIGGAFVDPKPSYMYISPEALENFCYDMKT 182
 DB 145 SMDDLSIONLGTKLASQMHKLTSLNLRIGGAFVDPKPSYMYISPEALENFCYDMKT 204
 QY 183 TCLPMFGYKHVLTLDQVTRFNEEVKQSVSRNRDAPEGGFDAIMQATVDEKIGMRNDA 242
 DB 205 TCLPMFGYKHVLTLDQVTRFNEEVKQSVSRNRDAPEGGFDAIMQATVDEKIGMRNDA 264
 QY 243 SHLVFTTDAKTHIALDGRLAGIVQPDGQCHVSDNHSASTTMDYPSLGLMTEKLSQK 302
 DB 265 SHLVFTTDAKTHIALDGRLAGIVQPDGQCHVSDNHSASTTMDYPSLGLMTEKLSQK 324
 QY 303 NINLIFAVTEENVNLXONYSELPFTTVGVLSMSSNVQLIYDAVGKIRSKVELEVRDL 362
 DB 325 NINLIFAVTEENVNLXONYSELPFTTVGVLSMSSNVQLIYDAVGKIRSKVELEVRDL 384
 QY 363 PEEILSLFNATCLNNEVTPGLKSCMGKLTIGDIVSFSIEAKVRGCPQEKESFTIKPVGFR 422
 DB 385 PEEILSLFNATCLNNEVTPGLKSCMGKLTIGDIVSFSIEAKVRGCPQEKESFTIKPVGFR 444
 QY 423 DSLIYQVTFDCCACQAOAEPNHRNNGNGTEFCGVCAGGWMLSQCECSEEDYRPSQ 482
 DB 445 DSLIYQVTFDCCACQAOAEPNHRNNGNGTEFCGVCAGGWMLSQCECSEEDYRPSQ 504

Db 445 DSIITQVTRDCCACAOAEPSSHRGNNGNTEFEGVCCLGPGWLGSCQECSEEDYHPSQ 504
 Qy 483 QDECSFREGQPVCSQREGELCGQCVCHSSDFGKITGKYCECDDFSCVRYKGMCSGHQC 542
 Db 505 QDECSFREGQPVCSQREGELCGQCVCHSSDFGKITGKYCECDDFSCVRYKGMCSGHQC 564
 Qy 543 SCGDCLCDSDDMTGYCNCCTTRDTGCMSSNGLLCSGRKCECCSCVCIOPGSYGDTCEKCP 602
 Db 565 SCGDCLCDSDDMTGYCNCCTTRDTGCMSSNGLLCSGRKCECCSCVCIOPGSYGDTCEKCP 624
 Qy 603 TCPDCTFKECKEYCECKEFGALHDENTCNRYCRDEIESVKELKDTGDAVACTYKNEED 662
 Db 625 TCPDCTFKECKEYCECKEFGALHDENTCNRYCRDEIESVKELKDTGDAVACTYKNEED 684
 Qy 663 CVVRFQYEDSSGKSTLYVEEPECKGPDILVLLSVGAILLIGLALLIMKLLITLTH 722
 Db 685 CVVRFQYEDSSGKSTLYVEEPECKGPDILVLLSVGAILLIGLALLIMKLLITLTH 744
 Qy 723 DRKEFAKFEERARAKWDTANNPLYKEATSTFTNTTYRGT 762
 Db 745 DRKEFAKFEERARAKWDTANNPLYKEATSTFTNTTYRGT 784

RESULT 3
 ID 09TUN3 PRELIMINARY: PRT: 788 AA.
 AC 09TUN3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GLYCOPROTEIN IIIA.
 GN GPIIa.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLATELET;
 RA Tao J., Parrilla R.;
 RT "Molecular cloning of rabbit platelet GPIIa cDNA."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: AF170529; AAD51955.1; -
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR Pfam: PR00362; Integrin_B.1.
 DR PRINTS: PRO1186; INTEGRINB.
 DR PRODOM: PD001811; Integrin_B.1.
 DR SMART: SM00187; INB.1.
 DR SMART: SM00423; PSI.1.
 DR SMART: SM00327; VWA.1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS. UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2.1.
 DR PROSITE: PS00243; INTEGRIN_BETA.3.
 DR Cell adhesion: Cytokeleton; EGF-like domain; Extracellular matrix;
 KM Glycoprotein; Integrin; Repeat; Transmembrane.
 SO SEQUENCE 788 AA; 87049 MW; 7ADB9CDN7301D78D CRC64;

Query Match 96.2%; Score 3996; DB 6; Length 788;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 726; Conservative 21; Mismatches 15; Indels 0; Gaps 0;
 Qy 1 GPNICTTRGVSSCQCCCLAVSPMCAMCSDEALPLGSPROCLKENLKDNCAPESIEFPVSE 60

Db 27 GPNICTTRGVSSCQCCCLAVSPMCAMCSDEALPLGSPROCLKENLKDNCAPESIEFPVSE 86
 Qy 61 ARVLEDRPLSKSGSGSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDL 120
 Db 87 AOLLEARPLSKSGSGSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDL 146
 Qy 121 SYSMKDDIWSIONLGYKLATOMRKLTSLNLRIGFAGYADKPVSPMYTSPPEALENCTYDM 180
 Db 147 SYSMKDDIWSIONLGYKLATOMRKLTSLNLRIGFAGYADKPVSPMYTSPPEALENCTYDM 206
 Qy 181 KTTCLPMFGKHYVLTLDQVTRFNEEYKOSVSRNRPDAEGGFATIMQATYCDKIGWRN 240
 Db 207 KTTCLPMFGKHYVLTLDQVTRFNEEYKOSVSRNRPDAEGGFATIMQATYCDKIGWRN 266
 Qy 241 DASHLVFTTDAKTHIALDGRAGIIVOPNDQCHVSGSDNHYASATMDYPSLGMTKLS 300
 Db 267 DASHLVFTTDAKTHIALDGRAGIIVOPNDQCHVSGSDNHYASATMDYPSLGMTKLS 326
 Qy 301 QKNINLFAVTENVVNLQYNSLIPGTTVGVLSMDSNNVQLLVDAVGIKRSVLEVR 360
 Db 327 QKNINLFAVTENVVNLQYNSLIPGTTVGVLSMDSNNVQLLVDAVGIKRSVLEVR 386
 Qy 361 DLPEELSLFNAATCLNNEVITPGLKSCMGKIGDVFSPFIEAKVNGCPQEKESFTIKPVG 420
 Db 387 DLPEELSLFNAATCLNNEVITPGLKSCMGKIGDVFSPFIEAKVNGCPQEKESFTIKPVG 446
 Qy 421 FKDSLIVQVTFDCCACAOAEPSSHRGNNGNTEFEGVCCLGPGWLGSCQECSEEDYR 480
 Db 447 FKDSLIVQVTFDCCACAOAEPSSHRGNNGNTEFEGVCCLGPGWLGSCQECSEEDYR 506
 Qy 481 SQDECSFREGQPVCSQREGELCGQCVCHSSDFGKITGKYCECDDFSCVRYKGMCSGHG 540
 Db 507 SQDECSFREGQPVCSQREGELCGQCVCHSSDFGKITGKYCECDDFSCVRYKGMCSGHG 566
 Qy 541 QCSGDCCLCDSDDMTGYCNCCTTRDTGCMSSNGLLCSGRKCECCSCVCIOPGSYGDTCEK 600
 Db 567 QCSGDCCLCDSDDMTGYCNCCTTRDTGCMSSNGLLCSGRKCECCSCVCIOPGSYGDTCEK 626
 Qy 601 CPTCPDCTFKECKEYCECKEFGALHDENTCNRYCRDEIESVKELKDTGDAVACTYKNE 660
 Db 627 CPTCPDCTFKECKEYCECKEFGALHDENTCNRYCRDEIESVKELKDTGDAVACTYKNE 686
 Qy 661 DDCVRFQYEDSSGKSTLYVEEPECKGPDILVLLSVGAILLIGLALLIMKLLIT 720
 Db 687 DDCVRFQYEDSSGKSTLYVEEPECKGPDILVLLSVGAILLIGLALLIMKLLIT 746
 Qy 721 IHDKEFAKFEERARAKWDTANNPLYKEATSTFTNTTYRGT 762
 Db 747 IHDKEFAKFEERARAKWDTANNPLYKEATSTFTNTTYRGT 788

RESULT 4
 ID 095JH1 PRELIMINARY: PRT: 784 AA.
 AC 095JH1;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GLYCOPROTEIN GPIIa.
 GN CD61.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ARTIC SMOOTH MUSCLE;
 RA Jimenez-Marin A.M., Garrido J.J., Iñanes D., Barhanchio M.J.;
 RT "Characterization of the porcine CD61 (GPIIa) gene."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF282890; AAK69529.1; -
 SO SEQUENCE 784 AA; 86365 MW; 1AE756651EEF3B8 CRC64;

Query Match 93.8%; Score 3898; DB 6; Length 784;
 Best Local Similarity 92.8%; Pred. No. 0;
 Matches 707; Conservative 35; Mismatches 20; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCCOQCLAVSPMCAMCSDDEALPLGSPRCDLKENILKDNCAPESEIEFPVSE 60
 DB 23 GPNICATRGVSSCCQCLAVSPMCAMCSDDEALPLGTPRCNLEENLNHCNPKYIEFPISE 82
 QY 61 ARVLEDRLPSDKSGSDSSQTVSPQRIALRLRDPDSKNFSIOVQVEDYVPVDIYYLMDL 120
 DB 83 ARILEARPLPSDKSGSDSSQTVSPQRIALRLRDPDSKNFSIOVQVEDYVPVDIYYLMDL 142
 QY 121 SYSKMDLMSIONLGTATLQATOMRKLTSNLRIGGAFYDKVSPMYTISPPEALENPPCYDM 180
 DB 143 SFSKMDLESIONLGTATLQASOMRKLTSNLRIGGAFYDKVSPMYTISPPEALENPPCYDM 202
 QY 181 KTTCLPMFGYKXHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRN 240
 DB 203 KTTCLPMFGYKXHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRN 262
 QY 241 DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHVSDNHSASTTMDYPSLGLMTERKLS 300
 DB 263 DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHVSDNHSASTTMDYPSLGLMTERKLS 322
 QY 301 OKNINLIFAVTVENVNLQVNSSELIPGTVGLMSDSSNVLOLIVDAVYGRIRSVLEVR 360
 DB 323 OKNINLIFAVTVENVNLQVNSSELIPGTVGLMSDSSNVLOLIVDAVYGRIRSVLEVR 382
 QY 361 DLPEELSLSFNATCLNNEVILPGKSCMGKLTIGTVSFSIEBAKVGCPQEKESFTITKPVG 420
 DB 383 DLPEELSLSFNATCLNNEVILPGKSCMGKLTIGTVSFSIEBAKVGCPQEKESFTITKPVG 442
 QY 421 FKOSLIQVTFEDDCAQQAEPNSHRCNNGNTEFEGVRCRCPMGLSGCESEEDYR 480
 DB 443 FKOSLIQVTFEDDCAQQAEPNSHRCNNGNTEFEGVRCRCPMGLSGCESEEDYR 502
 QY 481 SQODECSREGOPVCSQGRGECCLGQCCVCHSDGFKITGKCECDDEFCVYRKGEMSGHG 540
 DB 503 SQODECSREGOPVCSQGRGECCLGQCCVCHSDGFKITGKCECDDEFCVYRKGEMSGHG 562
 QY 541 QCSGDCGLCDSDMTGYCNCCTTRTDTGCMSSNGILCSGRGKCEGSCVCIQPGSYGDTCEK 600
 DB 563 QCSGDCGLCDSDMTGYCNCCTTRTDTGCMSSNGILCSGRGKCEGSCVCIQPGSYGDTCEK 622
 QY 601 CPTCPDACPTEKKEVECKKEDRGALHDENCNRYCRDELESYKELDTGDAANCCTYKNE 660
 DB 623 CPTCPDACPTEKKEVECKKEDRGALHDENCNRYCRDELESYKELDTGDAANCCTYKNE 682
 QY 661 DDCVVRFOYVEDSSGKSILVVEEPECPKGPDLVLLSVMAAILLIGLAAILMKLLLT 720
 DB 683 DDCVVRFOYVEDSSGKSILVVEEPECPKGPDLVLLSVMAAILLIGLAAILMKLLLT 742
 QY 721 IHRKKEFAKFEERARAKMDTANNPLKCATSTFTNITTYRGT 762
 DB 743 IHRKKEFAKFEERARAKMDTANNPLKCATSTFTNITTYRGT 784

RESULT 5
 Q9TUN5 PRELIMINARY; PRT; 784 AA.

AC Q9TUN5; 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE GLYCOPROTEIN IITA.
 GN GPITIA.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=PLATELET;
 RA Tao J., Parrilla R.;
 RT "Molecular cloning of pig platelet GPITIA cDNA";
 RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: AF170527; AAD51953.1;
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR002035; vWFA.
 DR Pfam: PF00362; Integrin_B; 1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 3.
 DR Cell adhesion; Cytoskeleton; EGF-like domain; Extracellular matrix;
 KW Glycoprotein; Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 784 AA; 86399 MW; EDED5651EE288D5 CRC64;

Query Match 93.7%; Score 3894; DB 6; Length 784;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 706; Conservative 35; Mismatches 21; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCCOQCLAVSPMCAMCSDDEALPLGSPRCDLKENILKDNCAPESEIEFPVSE 60
 DB 23 GPNICATRGVSSCCQCLAVSPMCAMCSDDEALPLGTPRCNLEENLNHCNPKYIEFPISE 82
 QY 61 ARVLEDRLPSDKSGSDSSQTVSPQRIALRLRDPDSKNFSIOVQVEDYVPVDIYYLMDL 120
 DB 83 ARILEARPLPSDKSGSDSSQTVSPQRIALRLRDPDSKNFSIOVQVEDYVPVDIYYLMDL 142
 QY 121 SYSKMDLMSIONLGTATLQATOMRKLTSNLRIGGAFYDKVSPMYTISPPEALENPPCYDM 180
 DB 143 SFSKMDLESIONLGTATLQASOMRKLTSNLRIGGAFYDKVSPMYTISPPEALENPPCYDM 202
 QY 181 KTTCLPMFGYKXHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRN 240
 DB 203 KTTCLPMFGYKXHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRN 262
 QY 241 DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHVSDNHSASTTMDYPSLGLMTERKLS 300
 DB 263 DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHVSDNHSASTTMDYPSLGLMTERKLS 322
 QY 301 OKNINLIFAVTVENVNLQVNSSELIPGTVGLMSDSSNVLOLIVDAVYGRIRSVLEVR 360
 DB 323 OKNINLIFAVTVENVNLQVNSSELIPGTVGLMSDSSNVLOLIVDAVYGRIRSVLEVR 382
 QY 361 DLPEELSLSFNATCLNNEVILPGKSCMGKLTIGTVSFSIEBAKVGCPQEKESFTITKPVG 420
 DB 383 DLPEELSLSFNATCLNNEVILPGKSCMGKLTIGTVSFSIEBAKVGCPQEKESFTITKPVG 442
 QY 421 FKOSLIQVTFEDDCAQQAEPNSHRCNNGNTEFEGVRCRCPMGLSGCESEEDYR 480
 DB 443 FKOSLIQVTFEDDCAQQAEPNSHRCNNGNTEFEGVRCRCPMGLSGCESEEDYR 502
 QY 481 SQODECSREGOPVCSQGRGECCLGQCCVCHSDGFKITGKCECDDEFCVYRKGEMSGHG 540
 DB 503 SQODECSREGOPVCSQGRGECCLGQCCVCHSDGFKITGKCECDDEFCVYRKGEMSGHG 562


```

QY 541 QCSGCDLSDSDMTGYCNCRTTRDTQMSNGLLSGRGKCEGCVCIQPSYGDTCER 600
DB 563 QCSGCDLSDSDMTGYCNCRTTRDTQMSNGLLSGRGKCEGCVCIQPSYGDTCER 622
QY 601 CPTCPDAGTFFKCCVCKCFDRCALHDEMTCNRYCDELESKELDKTKDQVANCYKNE 660
DB 623 CPTCPDAGTFFKCCVCKCFDRCALHDEMTCNRYCDELESKELDKTKDQVANCYKNE 682
QY 661 DDCVRFVRYEDSSGKSTLYVEEPCPKGPDLVLLVSMGAILLIGLALIMKLLT 720
DB 683 DDCVRFVRYEDSSGKSTLYVEEPCPKGPDLVLLVSMGAILLIGLALIMKLLT 742
QY 721 HDRKEFAKFEERARAKMDTANNPLYKEATSTFTNITYRGT 762
DB 743 HDRKEFAKFEERARAKMDTANNPLYKEATSTFTNITYRGT 784

RESULT 6
ID Q90W16 PRELIMINARY; PRT; 723 AA.
AC Q90W16;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DB BETA 3 INTEGRIN, GPIIA.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93290675; PubMed=8512576;
RA Cleutiaux A.M., Rosa J.P., Letourneur F., Poncez M., Rifat S.;
RT "A comparative analysis of cDNA-derived sequences for rat and mouse
RT beta 3 integrins (GPIIA) with their human counterpart.";
RC Biochem. Biophys. Res. Commun. 193:771-778(1993).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- PPM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
DR Interpro: IPR002086; Aldehyde_dehydr.
DR Interpro: IPR000561; EGF-like.
DR Interpro: IPR002369; Integrin_B.
DR Interpro: IPR001169; Integrin_beta-C.
DR Interpro: IPR002035; WMFA.
DR Pfam: PF00362; Integrin_B_1.
DR PRINTS: PR01186; INTEGRIN.
DR PRODOM: PD001811; Integrin_B_1.
DR SMART: SM00001; EGF-like; 1.
DR SMART: SM00187; INB; 1.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS00186; EGF_2; UNKNOWN_1.
DR PROSITE: PS00243; INTEGRIN_BETA_3.
KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
KW Integrin; Repeat; Transmembrane.
SQ SEQUENCE 723 AA; 80233 MW; B31D3127E6EC9F0F CRC64;

Query Match 86.88; Score 3606; DB 11; Length 723;
Best Local Similarity 90.58; Pred. No. 3e-291;
Matches 654; Conservative 42; Mismatches 27; Indels 0; Gaps 0;

```

```

DB 121 PVSPTMFIPTPOALKNPDTYTKSTCLPMEFGYKHYLTLDQVTKNEBKQOSVRNDAP 180
QY 220 EGGFPAIMQATVCEKIGMRNDASHLFTYTDATHTIALDRLAGIVQPDGQCHVSDN 279
DB 181 EGGFPAIMQATVCEKIGMRNDASHLFTYTDATHTIALDRLAGIVQPDGQCHVSDN 240
QY 280 HYSASTMDYPSLGLMTKLSOKNINLIFAVTEVNVLYONSELIPCTYGVUSMDSSN 339
DB 241 HYSASTMDYPSLGLMTKLSOKNINLIFAVTEVNVLYONSELIPCTYGVUSMDSSN 300
QY 340 VLQQLVDAVYGRKRSVLEVDLPDELSPNACLNNEVPGKSCGKLGIDTVPSFI 399
DB 301 VLQQLVDAVYGRKRSVLEVDLPDELSPNACLNNEVPGKSCGKLGIDTVPSFI 360
QY 400 EAKVRCPEQKEKESFTIRPVGFKDLSLYVTFDDCAQQAEPNSHRCNNGNTEFCGV 459
DB 361 EAKVRCPEQKEKESFTIRPVGFKDLSLYVTFDDCAQQAEPNSHRCNNGNTEFCGV 420
QY 460 CRGPGWLGSCGCESEEDYRPSQDECSFREGQPVCSRGCELCGQCCHSSDFGKTKG 519
DB 421 CRGPGWLGSCGCESEEDYRPSQDECSFREGQPVCSRGCELCGQCCHSSDFGKTKG 480
QY 520 YCECDDEFCVRYKGMCSGHGCGDCLDSDMTGYCNCRTTRDTQMSNGLLSGRG 579
DB 481 YCECDDEFCVRYKGMCSGHGCGDCLDSDMTGYCNCRTTRDTQMSNGLLSGRG 540
QY 580 KCEGSCVCIQPSYGDTCERKCPDCTEKKCECKEDRGALHDEMTCNRYCDEI 639
DB 541 KCEGSCVCIQPSYGDTCERKCPDCTEKKCECKEDRGALHDEMTCNRYCDEI 600
QY 640 ESKYELKDTGKRAVNCTYNNEDCVRRQYEDSSGKSTLYVEEPCPKGPDLVLLS 699
DB 601 ELVKELTDGKRAVNCTYNNEDCVRRQYEDSSGKSTLYVEEPCPKGPDLVLLS 660
QY 700 VMGAILLIGLALIMKLLITIHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNITY 759
DB 661 VMGAILLIGLALIMKLLITIHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNITY 720

QY 760 RGT 762
DB 721 RGT 723

RESULT 7
ID Q92071 PRELIMINARY; PRT; 781 AA.
AC Q92071; Q90829;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INTEGRIN BETA3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN;
RX MEDLINE=94164000; PubMed=8119143;
RA Miyaura H., Cao X.P., Ross F.P., Chiba M., Teitelbaum S.L.;
RT "1,25-dihydroxyvitamin D3 transcriptionally activates the beta 3-
RT integrin subunit gene in avian osteoclast precursors.";
RL Endocrinology 134:1061-1066(1994).
RN [2]
RP SEQUENCE OF 1-26 FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=94086557; PubMed=8262978;
RA Cao X., Ross F.P., Zhang L., MacDonald P.N., Chappel J.,
RA Teitelbaum S.L., Patrick F.;
RT "Cloning of the promoter for the avian integrin beta 3 subunit gene
RT and its regulation by 1,25-dihydroxyvitamin D3.";

```

RL J. Biol. Chem. 268:27371-27380(1993).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: X72378; CAAS1069.1; -
 DR EMBL: X75348; CAAS3095.1; -
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR pfam: PF00362; Integrin_B_1.
 DR PRINTS: PR01186; INTEGRINB.
 DR PRODOM: PD00181; Integrin_B_1.
 DR SMART: SM00001; EGF_Like_1.
 DR SMART: SM00187; INB_1.
 DR SMART: SM00423; PSI_1.
 DR SMART: SM00327; VWA_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 781 AA; 86115 MW; 7E991B605DFCDBC CRC64;

Query Match 83.9%; Score 3484.5; DB 13; Length 781;
 Best Local Similarity 83.0%; Pred. No. 4.4e-281;
 Matches 632; Conservative 58; Mismatches 70; Indels 1; Gaps 1;

QY 1 GNICTTGVSSCOQCLAVSPWCAMCSDALPLGSPROCKLKNLKDCAPEISFPSE 60
 DB 20 GSNICATRGVTSCKKCLAVSPICAMCSAEVVAQSTPRCDLFANLLONGCGRFIFPPSS 79
 QY 61 ARVLEPRPLSDKSGSDSSQVTVSQRALRLRPDCKNFSIQVQVEDYPVDIYYLMDL 120
 DB 80 IVLEERPLSDKSGSGST-TTQMSFORQLNLRPDQSMFRVHVQVEDYPVDIYYLMDL 138
 QY 121 SYSKMDLWASIONLGTATKATQMRKLTSLNRIGFAFVDPKPSPVYATISPPALENPCDM 180
 DB 139 SSMMDLKNIONLGTATKATQMRKLTSLNRIGFAFVDPKPSPVYATISPPALENPCDM 198
 QY 181 KTTCLPMFGYKHVLTLDQVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRN 240
 DB 199 GEKCLPMFGYKHVLTLDQVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRN 258
 QY 241 DASHLLVETPAKTHIALDGRLAGIVQNDGQCHVGSNNHYSASTTMDYPSLGLMTERKLS 300
 DB 259 DASHLLVETPAKTHIALDGRLAGIVQNDGQCHVGSNNHYSASTTMDYPSLGLMTERKLS 318
 QY 301 QKNINILFAVTENVVNLKQNSSELIPGTVGLSDSSNNVQLYDAVAGKIRSVLEEVLR 360
 DB 319 QKNINILFAVTENVVNLKQNSSELIPGTVGLSDSSNNVQLYDAVAGKIRSVLEEVLR 378
 QY 361 DLPEELSLSPNATCLNNEVIFGLKSCMGLKIGDYVFSIEAKVNGCQDEKESFTIPVVG 420
 DB 379 DLPEELSLSPNATCLNNEVIFGLKSCMGLKIGDYVFSIEAKVNGCQDEKESFTIPVVG 438
 QY 421 FKDSLIVQTPDCCACQAQAEPPNSHRCNNNGNTEFEGCVRCGPGMTLGSQCESEEDYRP 480
 DB 439 FKDSLIVQTPDCCACQAQAEPPNSHRCNNNGNTEFEGCVRCGPGMTLGSQCESEEDYRP 498
 QY 481 SQODECSPPREGQPYVCSQREGCELCGQCVCHSSDFKIGKVCCEODDFGCVRRKKGMSGSHG 540
 DB 499 SQODECSPPREGQPYVCSQREGCELCGQCVCHSSDFKIGKVCCEODDFGCVRRKKGMSGSHG 558
 QY 541 QCSGCDCLCDSQDWTGYCCNCTTRTDTCMSNGLICSGKCGKCGVCCVCIOPGSGVDTCEK 600
 DB 559 QCSGCDCLCDSQDWTGYCCNCTTRTDTCMSNGLICSGKCGKCGVCCVCIOPGSGVDTCEK 618
 QY 601 CPTCPDACTFKKECECKKFDKRGALHDEMTCNRYCRDEIESVKELKDTGKDAVANCYKNE 660

DB 619 CPTCPDACTFKKECECKKFERGTLEEQSGFRVCRDEIETVQELDGRKDAVANCYKNE 678
 QY 661 DDCVRFQVYEDSSGKSLTYVEEPCRGDIIIVLISWCAITLLGLALLKLLIT 720
 DB 679 NDCVRFQVYEDSSGKSLTYVEEPCRGDIIIVLISWCAITLLGLALLKLLIT 738
 QY 721 IHDRREFARFEERARAKMDTANPLYEATSTFTNITYRG 761
 DB 739 IHDRREFARFEERARAKMDTANPLYEATSTFTNITYRG 779

RESULT 8
 ID 09QW15 PRELIMINARY; PRT; 680 AA.
 AC 09QW15.
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE BETA 3 INTEGRIN, GPIIb.
 OS Mus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_Taxid=10095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93290675; PubMed-8512576;
 RA Jentat A.M., Rosa J.P., Letourneur F., Poncez M., Rifat S.;
 RT "A comparative analysis of cDNA-derived sequences for rat and mouse
 beta 3 integrins (GPIIb) with their human counterpart."
 RL Biochem. Biophys. Res. Commun. 193:771-778(1993).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR002035; VWA.
 DR Pfam: PF00362; Integrin_B_1.
 DR PRINTS: PR01186; INTEGRINB.
 DR PRODOM: PD00181; Integrin_B_1.
 DR SMART: SM00001; EGF_Like_1.
 DR SMART: SM00187; INB_1.
 DR SMART: SM00327; VWA_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 680 AA; 75514 MW; 672DB7338DD86003 CRC64;

Query Match 83.6%; Score 3472; DB 11; Length 680;
 Best Local Similarity 92.4%; Pred. No. 3.9e-280;
 Matches 628; Conservative 32; Mismatches 20; Indels 0; Gaps 0;

QY 79 QVTQVSFORIALRLRPDCKNFSIQVQVEDYPVDIYYLMDL.SYSKMDLWASIONLGTKL 138
 DB 1 QVTQVSFORIALRLRPDCKNFSIQVQVEDYPVDIYYLMDL.SYSKMDLWASIONLGTKL 60
 QY 139 ATQMRKLTSLNRIGFAFVDPKPSPVYATISPPALENPCYDMKTTCLPMFGYKHVLTLD 198
 DB 61 ASQMRKLTSLNRIGFAFVDPKPSPVYATISPPALENPCYDMKTTCLPMFGYKHVLTLD 120
 QY 199 QVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCEKIGWRNDASHLLVETPAKTHIAL 258
 DB 121 QVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCEKIGWRNDASHLLVETPAKTHIAL 180
 QY 259 DGRLAGIVQNDGQCHVGSNNHYSASTTMDYPSLGLMTERKLSQKNINILFAVTENVVNL 318
 DB 318 DGRLAGIVQNDGQCHVGSNNHYSASTTMDYPSLGLMTERKLSQKNINILFAVTENVVNL 318

```

Db 181 DGRILGIVLPNDGCHIGTDNHYASTMDYSLGLMTEKLSOKNINLFAVTENVSLY 240
QY 319 QNSELIPGTTVGLMSDSSNVQLLVDAVGRKRSKVELEVRDPELSLSPATCLNNE 378
Db 241 QNSELIPGTTVGLMSDSSNVQLLVDAVGRKRSKVELEVRDPELSLSPATCLNNE 300
QY 379 VIRGLKSCVGLKIGDVFSEIEAKVRCGPOEKESFTIKPVGFKDSLIVQVFPDCCACQ 438
Db 301 VIRGLKSCVGLKIGDVFSEIEAKVRCGPOEKESFTIKPVGFKDSLIVQVFPDCCACQ 360
QY 439 AQAEPSNRHCNNGNFTFEGCVCCGCGMGLASOCESEEDYRPSQODECSFREGQPVCSQR 498
Db 361 AFAPSSPRCNCNNGFTFEGCVCCGCGMGLASOCESEEDYRPSQODECSFREGQPVCSQR 420
QY 499 GECGCGCVCHSSDFGKITGKCECDDEFCVRYKKGEMSGHGCCSGDCLCDSDMTGYCC 558
Db 421 GECGCGCVCHSSDFGKITGKCECDDEFCVRYKKGEMSGHGCCSGDCLCDSDMTGYCC 480
QY 559 NCTTRDTGCMSSNGLCSGRKCEGSCVCIOPGSGYDTCCEKPCPDACTEFKCEVECK 618
Db 481 NCTTRDTGCMSSNGLCSGRKCEGSCVCIOPGSGYDTCCEKPCPDACTEFKCEVECK 540
QY 619 KEDRGALHDENTCNRVCRDEIESYKELDKTGKDAVNTYKNEDDCVYRFGYEDSSGSKI 678
Db 541 KFNRTGHEERTCSRYCRDIEQYKELDTGKNAVNTYKNEDDCVYRFGYEDTSGRAY 600
QY 679 LYVEEPCRCRGPDIIVLVYMGAILLIGLALLIKLITTIHDKREFAKFEERARAK 738
Db 601 LYVEEPCRCRGPDIIVLVYMGAILLIGLALLIKLITTIHDKREFAKFEERARAK 660
QY 739 WDTANNPVLYKEATSTFTNIT 758
Db 661 WDTANNPVLYKEATSTFTNIT 680

RESULT 9
007012 PRELIMINARY: PRT: 788 AA.
ID 007012:
AC 007012:
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INTEGRIN BETA-3 SUBUNIT PRECURSOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC xenopodidae; Xenopus.
CX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94040374; PubMed=7693527;
RA Ransom D.G., Hens M.D., Desimone D.W.;
RT "Integrin expression in early amphibian embryos: cDNA cloning and
RT characterization of Xenopus beta 1, beta 2, beta 3, and beta 6
RT subunits.";
RL Dev. Biol. 160:265-275(1993).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
DR EMBL, L13591: AA17427.1;
DR InterPro: IPR002086; AIdenhydr_dehydr.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002369; Integrin_B.
DR InterPro: IPR001169; Integrin_beta_C.
DR InterPro: IPR003659; PSI.
DR Pfam: PF00362; Integrin_B_1.
DR PRINTS: PR01186; INTEGRINB.
DR ProDom: PD001811; Integrin_B_1.
DR SMART: SM00181; EGF_1.
DR SMART: SM00187; INB_1.
DR SMART: SM00423; PSI_1.
DR PROSITE: PS00070; ALDENHDE_DEHYDR_CYS; UNKNOWN_1.

```

```

DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS00243; INTEGRIN_BETA_2.
KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
KW Integrin; Repeat; Signal; Transmembrane.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 788 INTEGRIN BETA-3 SUBUNIT.
SQ SEQUENCE 788 AA: 87490 MW: 80902510E24A6109 CRC64:

Query Match 78.5%; Score 3260.5; DB 13; Length 788;
Best Local Similarity 75.9%; Pred. No. 1.9e-262;
Matches 578; Conservative 89; Mismatches 94; Indels 1; Gaps 1;

QY 1 GPNICTTRGVSSCOQCLAVSPMCANCSDEALPLGSPRCDLKENLKDNCAPESIEFPYSE 60
Db 24 GASICATRCVSSCOQCLAVSPMCANCSDEALPLGSPRCDLKENLKDNCAPESIEFPYSE 83
QY 61 ARVLEDRPLSDKGS-GDSQVQTVSPORALRLRPDSDKNSIOVROYEDYPVDTIYIMD 119
Db 84 VLVAEHRPLSVKSGSDREITQMSPOKIDFLRPDDTKVFMLOVROYEDYPVDTIYIMD 143
QY 120 LSYMKDGLMSIONLGTATQMRKLTSLNLRIGFAFYDKPVSPYMYISPPALENPGYD 179
Db 144 LSYMKDGLIKIQTLSLSERMRITSLNLRIGFAFYDKPVSPYMYISPPALENPGYD 203
QY 180 MTTCLPMFGYKHVLTLDQVTRFNEEYKQSVSRNRDAPESGFALINQATYCDKIGMR 239
Db 204 FTECPMPTEGYKHVLTLEVRLEFNEEYQKQVSRNRDAPESGFALINQATYCDKIGMR 263
QY 240 NDASHILVFTTQAKHIALDGRLAGIVOPNDGCHVSNHYSASTNDYPSLGLMTEKL 299
Db 264 NESHLVFTTQAKHIALDGRLAGIVOPNDGCHVSNHYSASTNDYPSLGLMTEKL 323
QY 300 SOKNINLFAVTENVVLYKONTSELIPGTTVGLMSDSSNVQLLVDAVGRKRSKVELEV 359
Db 324 SOKNINLFAVTENVVLYKONTSELIPGTTVGLMSDSSNVQLLVDAVGRKRSKVELEV 383
QY 360 RDLPEELSLSFNATCLNNEVYIPGLKSCMGLKIGDYVSEIEAKVRCGPOEKESFTIKPV 419
Db 384 RDLPEELSLSFNATCLNNEVYIPGLKSCMGLKIGDYVSEIEAKVRCGPOEKESFTIKPV 443
QY 420 GFQDSLIVQVTPPCDCAQOAPNPNHRCNNGNFTFEGCVCCGCGMGLASOCESEEDYR 479
Db 444 GFQDSLIVQVTPPCDCAQOAPNPNHRCNNGNFTFEGCVCCGCGMGLASOCESEEDYR 503
QY 480 PSQODECSFREGQPVCSQRDEICLGQCVCHSSDFGKITGKCECDDEFCVRYKKGEMSGH 539
Db 504 PSQODECSFREGQPVCSQRDEICLGQCVCHSSDFGKITGKCECDDEFCVRYKKGEMSGH 563
QY 540 GCGSCGDCCLDSMTGYCCNCTRTDTCMSSNGLCSGRKCEGSCVCIOPGSGYDTCCE 599
Db 564 GCGSCGDCCLDSMTGYCCNCTRTDTCMSSNGLCSGRKCEGSCVCIOPGSGYDTCCE 623
QY 600 KCPTCPDACPTEFKCEVECKKFDGALHDENTCNRVCRDEIESYKELDKTGKDAVNTYRN 659
Db 624 KCPTCPDACPTEFKCEVECKKFDGALHDENTCNRVCRDEIESYKELDKTGKDAVNTYRN 683
QY 660 EDDCVARFQYEDSSGSKILYVEEPCRCRGPDIIVLVYMGAILLIGLALLIKLIT 719
Db 684 EDDCVARFQYEDSSGSKILYVEEPCRCRGPDIIVLVYMGAILLIGLALLIKLIT 743
QY 720 TIHDKREFAKFEERARAKWDTANNPVLYKEATSTFTNITIRG 761
Db 744 TIHDKREFAKFEERARAKWDTANNPVLYKEATSTFTNITIRG 785

RESULT 10
09GK49 PRELIMINARY: PRT: 791 AA.
AC 09GK49:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

```

DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE INTERLIN BETA-5 SUBUNIT (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 NCBI_Taxid=9913;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY GLAND;
 RA Andersen M.H., Rasmussen J.T., Berglund L., Petersen T.E.;
 RT Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF317198; AAG36594.1;
 DR EMBL: AF317198; AAG36594.1;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR003659; PSI.
 DR Pfam: PF00362; Integrin_B.
 DR Pfam: PF00362; Integrin_B.
 DR PRINTS: PRO1186; INTEGRINB.
 DR ProDom: PD00181; Integrin_B; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00001; EGF_like; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN.1.
 DR NON_TER 1
 SQ SEQUENCE 791 AA; 87144 MW; 3FB45E62374169BC CRC64;

Query Match 55.7%; Score 2315.5; DB 6; Length 791;
 Best Local Similarity 54.8%; Pred. No. 8.4e-184;
 Matches 426; Conservative 122; Mismatches 211; Indels 19; Gaps 9;

1 GENICTRGSSCOOCLAVSPWCAMCSD---ALPLGSPRODLKKNLKDNCAPSIIEP 57
 15 GINICTSGATSCCECLIHPCACWCFEDGSLSVSRCDLKNLKNCCGVE-FESP 73
 58 VSEARLEDRPLSDKSGDS--SQVTSVPORIALRPPDSKNFSIQROYEDYVVDIY 116
 74 ASSTQVLSLPLSSKSGSPAGSDVQLPQEVTVTLRPGDTRATQLOQROYEDYVVDIY 133
 117 LMDLSYSKDDLSIQNTGKTATQMKLTSLNLRIGAFYVDKVPYMYISPEALENP 176
 134 LMDLSLSKDDLDNIRSLGTIKLAEMRKLTNFRIGFSGVDKNTSPSY-TAPRYGTP 192
 177 C--YDMKTTCLPMFGYKRVLTLDQVTRFNEEVKKQSVSRNRDAPEGFDIMATYCD 234
 193 CIGYLPNCVPSFGFRHLPLTDVDSFNEEVKQSVSRNRDAPEGFDVLAQAAVCKE 252
 235 KIGWRNDASHLLVFTTDAKTHIALDGLAGIVQPNDCGCHVSGSNHYSASTTMDYPSLGL 294
 253 KIGWRKDALHLVFTTDOVPHIALDGLKGLVQPHDQCHLNEANETASNDMPYSLAL 312
 295 MTEKLSQNTLIFAVENYVNLQNTSELIPGTVGLMSDSSNVLLQILVDAYGKTRSK 354
 313 LGEEKLAENNNILIFAVYKNHMYLKNFTALIPGTVGLMSDSSNVLLQILINAVYSIRSK 372
 355 VELEVRDPEELSLFNATCLNNEVIPGKSKMGKLTGDMYPSFIEAKVGGCQPEK-EKS 413
 373 VELSVMPQEDLNLFATATCODGYSTGQKRCBGLKIGDTRASLRVSEARSCPSKHEHT 432
 414 FTIKPVGRKSLIVQVTFDDCCAOAOAPNSHRNNGNGSTFECCVRCRGPWGLSQCEC 473
 433 LALRPVAFRDSLEGVYTYNRCGSGAGLEPDSARCS-NGTYVCGGLECPMGLYGTCEC 491
 474 SEELRYASQDECSRGQVPCSRGCLIGQCYCHSDGKLTGKTCEDDPSCVARYK 533
 492 QEGSGSGYQMLCREAGKPKLCSGQSCNOCSCFESSEPKIYGSFCEDNFCARANK 551
 534 EMCSGHQSCGDLCDSDMTGYCNCCTTRDPCMSNGGLSGRGKCEGSCVCTOPGS 593
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

DB 552 VLCSGHGECHECKCHAGYIGDNCNSTDSTQCARCGHICSDRGHCVCQCCCTEPCA 611
 QY 594 YDTCCEKPTCPDACEKFEKCECKECPDRGALHDENTONRCYRDE-IESVELMDTGDA 652
 DB 612 FGETCEKPTCPDACEKFEKCECKECPDRGALHDENTONRCYRDE-IESVELMDTGDA 671
 QY 653 VNCYKKNEDDCVRFQYVEDSSGKILVVEPECPKPDILLVLLSYMAHLLIGLAL 712
 DB 672 VLCSGHGECHECKCHAGYIGDNCNSTDSTQCARCGHICSDRGHCVCQCCCTEPCA 731
 QY 713 LMKLLITIHDRKEFAKFEERARAKMDTANPNLYKETSFT-----NITRYGT 762
 DB 722 YIMKLVITIHDRKEFAKFEERARAKMDTANPNLYKETSFT-----NITRYGT 789

RESULT 11
 ID O9GLP0 PRELIMINARY; PRT; 798 AA.
 AC O9GLP0;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE INTEGRIN BETA-1 SUBUNIT.
 GN CD29.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_Taxid=9623;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20426040; Pubmed-10972224;
 RA Jimenez-Marin A., Garrido J.J., de Andres-Cara D.F., Moreta L.,
 RA Barbancho M.J., Llanes D.;
 RT "Molecular cloning and characterization of the pig homologue to human
 RT CD29, the integrin beta1 subunit."
 RL Transplantation 70:649-655(2000).
 DR EMBL: AF192528; AAG16767.1;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR002035; VWA.
 DR Pfam: PF00362; Integrin_B; 1.
 DR PRINTS: PRO1186; INTEGRINB.
 DR ProDom: PD00181; Integrin_B; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN.1.
 SQ SEQUENCE 798 AA; 88265 MW; 80773E985E38BA90 CRC64;

Query Match 44.1%; Score 1833; DB 6; Length 798;
 Best Local Similarity 44.2%; Pred. No. 1.2e-143;
 Matches 345; Conservative 140; Mismatches 267; Indels 28; Gaps 13;

3 NITTTGSSCOOCLAVSPWCAMCSD---EALPLGSPRODLKKNLKDNCAPSIIEP 57
 25 NRCILNAKSGAGCICDAGPNCGMCNTSTLQSGMP-SARCDLLELRKKGCHPDIDNP 83
 58 VSEARLEDRPLSDKSGDSQ---YVQTSVPORIALRPPDSKNFSIQROYEDYVVDIY 112
 84 RGSKNIKKNNKNTNRSKGAEKLPEDIDQIQPQOLVLDLRSGEQOTFLTKKRAEDYPI 143
 113 DLYIMDLISYKDDLSIQNTGKTATQMKLTSLNLRIGAFYVDKVPYMYISPEEA 172
 144 DLYIMDLISYKDDLSIQNTGKTATQMKLTSLNLRIGAFYVDKVPYMYISPEEA 202
 173 LENPCDMKTTCLPMFGYKRVLTLDQVTRFNEEVKKQSVSRNRDAPEGFDIMATYCD 232
 203 LKRPC-TSQONTSPFSTKYNVSLTDKGEVVELVKORISLNDSPGCGDPAIMQVAVC 261
 233 DEKIGWRNDASHLLVFTTDAKTHIALDGLAGIVQPNDCGCHVSGSNHYSASTTMDYPSL 292

```

DB 262 GSLIGRNR-VTRLLVSTDAAGFFACDGLGVLDPNDGCHLEND-VYTMSSHYVYPSI 319
QY 293 GLMTERKLSOKNINLIFAVTENVNLYONTSSELIPGTTGVLSDDSNVLIQIYDAYGKIR 352
DB 320 AHLVOKLSNNIOTIFAVTEEPQVYKELKNIIPKSAVGTLSANSSNVLIQIYDAYGKIR 379
QY 353 SKYELVRDLPBELSLSFNATCLN-NEVYIPGLSKSMGLKIGTYSFSEIAYRGCPQEK 410
DB 380 SEVILENSKLPBEVTTINYSYCKNNGVGENGRKCSNISTGDEVQFELISTANKCPKNK 439
QY 411 EKSFITIPVGEKSLIVQVTFDDCAQOAPBNSHRCNNGNGTFECGVCRCGPGMLGSO 470
DB 440 SEITIKIPPLGFTBEVEILLIOTICECEQSGEIPSPKCHDNGNTPFCGACRCNEGRVGRH 499
QY 471 CESEEDYRPSQODE-CPSPREGQVCSQRECELCGQCVCHSSDPGR--ITGKCECDNDS 527
DB 500 CEESTDEVSNDMDAYCKRNSSEICTNNGECVCGOCVCKRRDNTNEIYSGKFCEDDNEN 559
QY 528 CVRYKGMCSGHGCGSCGDCDMDTGYCNCCTRTDTGMSNGLLSGRGKCEGSCV 587
DB 560 CDSNGLICGNGYCKRVCBCNNTGSGACDSDLTSCMAVNGQICNGRGVCEGVCYK 619
QY 588 CIOPGSYDTCCEKPCPDACFEKCEVECKKFDGALHDENTCNRYCR---DELESVK 643
DB 620 CTDPKFGPGTCMGCOTCLGCAEHEKCEVQCRAFKNGEKD--TCAQCSHFHNTKVENRD 677
QY 644 ELKRTGR--DAVNCTYKNEDDCVVRFOYEDSSGKSLTYVEEPECKPGDIIIVLLSV 701
DB 678 KLPQGVGDPDLSCHEKEDVDCCWFYFYSVNGNEAVHVEHTEPCTGDIIPYAGV 737
QY 702 GAILLIGLALLIKLITITIDRKEFAKFEERARAKMDTANPLYEATSTFTNTITYRG 761
DB 738 AGYVLIGLALLIKLIMITHDRERFAKFEKKNAMKMDGEMNIYSATVTYVNPKEG 797

RESULT 12
Q91A01 PRELIMINARY; PRT; 807 AA.
AC 091A01:
DB 01-OCT-2000 (TREMblrel, 15, Created)
DB 01-OCT-2000 (TREMblrel, 15, Last sequence update)
DB 01-DEC-2001 (TREMblrel, 19, Last annotation update)
DB BETA-1 INTEGRIN.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Siluriformes; Ictaluridae; Ictalurus.
OC NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RA Almsworth J.A., Qian Y., Noya M.;
RT "Molecular characterization and leukocyte distribution of teleost
RT beta-1 integrin molecule";
RL Submitted (JAN-2000) to the EMBL/GenBank/CDR databases.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- PFM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
DB EMBL: AF224337; AAF35883.1;
DR InterPro: IPR002086; Aldhyde_dehydr.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002369; Integrin_B.
DR InterPro: IPR001169; Integrin_beta_C.
DR InterPro: IPR003659; PSI.
DR InterPro: IPR002035; WMPA.
DR Pfam: PF00362; Integrin_B_1.
DR PRINTS: PR01186; INTEGRINB.
DR ProDom: PD001811; Integrin_B_1.
DR SMART: SM00001; EGF_like; 1.
DR SMART: SM00187; INB; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00327; VWA; 1.

```

```

DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00223; EGF_1; UNKNOWN_2.
DR PROSITE: PS00243; INTEGRIN_BETA; 3.
DR Cell adhesion: Cytoskeleton; Extracellular matrix; Glycoprotein;
KW Integrin; Repeat; Transmembrane.
SQ SEQUENCE 807 AA; 89368 MW; D7A30FE92D82B00C CRC64;

```

```

Query Match 43.3%; Score 1800.5; DB 13; Length 807;
Best Local Similarity 44.2%; Pred. No. 6.2e-141;
Matches 346; Conservative 132; Mismatches 270; Indels 35; Gaps 15;

```

```

QY 3 NICTTRGVSSCOCLAVSPBMCAMCSDEA-LPLG--SPRCDLEKLLDNCAPESIEEPV 58
DB 35 NDCITRASQSCGCEIOVGEKCGMCTYESFLKQGESKSTRCEJTEALLIKGCPPTNENPR 94
QY 59 SEARVLEDPRLSDKSGS-----DSOVQVSPORALRLRPDSKNFSIOVROVEDYVD 113
DB 95 GTIVANENKTVTRNKDVYKELKPDQITQIQOKLTLNLRAGEPQKFLKRRADYDID 154
QY 114 IYVLDLSYSMDLMSIONLGTKLATQMRKLTSLRIGFAYDKPVSPWYISPPAL 173
DB 155 LYVLDMSYSMDLENVNKLGTDLREMQSVTSDFRIGGFSFVETWMPYISTPAKLL 214
QY 174 ENPCIDAKTTCIPMGYKAVLTITDOVTRFNEVYKQSVSRKRDAPCEGPDALMGATVCD 233
DB 215 -NPC-TSDONCTSPSYKKNVLSLTEDGRQFNSLVSRQOISGMLJSDPEGFDIOMAVCG 272
QY 234 EKIGRNDASHLVFTTDAKTHIALDGRLAGIOPNDGCHGSDNHYSASTYMDYPSLG 293
DB 273 EQLGRNR-VTRLLVSTDAAGFFACDGLGVLDPNDGCHL-KNLLTMSHYVYPSIA 330
QY 294 LMTKLSQKNINLIFAVTENVNLYONTSSELIPGTTGVLSDDSNVLIQIYDAYGKIR 353
DB 331 HLYQKLSDDNNIOTIFAVTEEPQVYKELKNIIPKSAVGTLSANSSNVLIQIYDAYGKIR 390
QY 354 KYVLEVRDLPBELSLSFNATCLN-NEVYIPGLSKSMGLKIGTYSFSEIAYRGCPQEK 411
DB 391 EYILENSKLPBEVTTINYSYCKNNGVGENGRKCSNISTGDEVQFELISTANKCPKNK 450
QY 412 KSFTIPVGEKSLIVQVTFDDCAQOAPBNSHRCNNGNGTFECGVCRCGPGMLGSO 471
DB 451 SEVKIKIPPLGFTBEVEILLIOTICECEQSGEIPSPKCHDNGNTPFCGACRCNEGRVGRH 510
QY 472 ECSEEDYRPSQODE-CPSPREGQVCSQRECELCGQCVCHSSD--FEKIGKCECDNDS 528
DB 511 ECKSDVRAEDIDANCRMDGDTICSNNGDCTGCECKKRDPEKRYGKICECDNFC 570
QY 529 VRYKGMCSGHGCGSCGDCDMDTGYCNCCTRTDTGMSNGLLSGRGKCEGSCV 588
DB 571 DRSNKLICGNGYCKRVCBCNNTGSGACDSDLTSCMAVNGQICNGRGVCEGVCYK 619
QY 589 IOPGSYDTCCEKPCPDACFEKCEVECKKFDGALHD-ENTCVRVCRDELESVKELK 646
DB 631 TDPKFGPGTCMGCOTCLGCAEHEKCEVQCRAFKNGEKD--TCAQCSHFHNTKVENRD 677
QY 647 DTGKD-----AVNCTYKNEDDCVVRFOYEDSSGKSLTYVEEPECKPGDIIIVLL 698
DB 688 DLQPNKSKPYLSCKEKDANDCCWFYFYSVNGNEAVHVEHTEPCTGDIIPYAGV 737
QY 699 SYMGAILLIGLALLIKLITITIDRKEFAKFEERARAKMDTANPLYEATSTFTNT 758
DB 744 GLVAGYVLIGLALLIKLIMITHDRERFAKFEKKNAMKMDGEMNIYSATVTYVNPKE 803
QY 759 YRG 761
DB 804 YEG 806

RESULT 13
Q96444 PRELIMINARY; PRT; 788 AA.
AC 096444:

```

DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE BETA INTEGRIN SUBUNIT.
 OS Biomphalaria glabrata (Bloodfluke planorb).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 CC Planorbidae; Biomphalaria.
 NCBI_TaxID=6526;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99173885; PubMed=10072774;
 RA Davids B.J., Wu X.J., Yoshino T.P.;
 RT "Cloning of a beta integrin subunit cDNA from an embryonic cell line
 RT derived from the freshwater mollusc, Biomphalaria glabrata."
 RL Gene 228:213-223(1999).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: AF060203; AAC67503.1; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR002049; Laminin_EGF.
 DR Pfam: PR00362; Integrin_B_1.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B_1.
 DR SMART: SM00001; EGF_like.1.
 DR SMART: SM00187; INB.1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 788 AA; 87632 MW; 3FE3DFAFE0848163 CRC64;

Query Match 39.9%; Score 1658; DB 5; Length 788;
 Best Local Similarity 42.7%; Pred. No. 4.3e-129;

Matches 329; Conservative 136; Mismatches 267; Indels 38; Gaps 20;

QY 12 SCQCCLAIVSPMCWCSDEALPGLSP-RCDIKENLLKD-NCAPESIEFVSEKRVLEDPRL 69
 DB 32 SCASCIATSKCEKCAVCVSYEDENLRCDTHENHRLGILYCPGDIQFPTDEVEKLKNO-- 89
 QY 70 SDRSGSDSQ-VTQVSPQRIALRLRDPDSKNFSIQVQVEDYVPIYIMLSYSMKDDL 128
 DB 90 -DVGGDNDPEDAVQVQPKRIKIRPNKPEVKILFRGAENYVDLLIFMLDSMEDDK 148
 QY 129 WSIQWLGRKLATQMRKLTSLNIGGAFVYKVPYMTISPEALENDCYDMKT---TC 184
 DB 149 EKLALLGRKIAEQMSAIRKRNRLGFSFVDVVSPPY-STVPOKIKMPC--KIYNGEPC 204
 QY 185 LPMFGYKVLTLTDQVTFENEBEVKQSVSRNRDAPEGGEFALIMATVDEKIGRRNDASH 244
 DB 205 EAPYEFKNQSLIDLETTFSQKVKARSGVLDAPEGGFDAIMQVACEDIEDIGRRPSRR 264
 QY 245 LLLVPTTHALDGLAGIIVQPNQGHVSDNNHVSATTTMDYPSLGLMTEKLSOKNI 304
 DB 265 MLVSTDAGFHAGDGKIGIIVTPNDGCHL-KNNLYSESSNLDYPSVSOJANKIKKRSV 323
 QY 305 NLITAVTENNVNLYQNSSELLPGTTVGVLSMDSNVQIOLYDAGKIRKVELEVRDLPE 364
 DB 324 SVITAVVDLPDIYEKLSKYESTTGLANDSSNIYTLIDNKKITSKVTLKADLIGE 383
 QY 365 ELISFNATCLNNEVITGLKSCMLKIGDYVSFSIEAKVRCPOEKE--KSFTIKPVGF 421
 DB 384 NITYDFSRGRGGEIK-TNCGSLKIGQSVFPAEVLALACPDKKKWLKERSIRPLGY 442
 QY 422 KDSLIIVTPECCACQ--AQAEFNSHRONGNGTFFECGVRCGPGWLGSCCECSEEDYR 479
 DB 443 QEKLTVELEIMICECCENAEENEILNSDKSNGNGTFFECGKSCSHPRGYGFCECKADDLT 502

QY 480 P8QD-DECSPRGQPGVSGRGECLGGCVCH--SSDFPK-ITGKYCEDDFSCVRYKREM 535
 DB 503 SKDSIKQCAPAPATLPCSRGSGVCGECLCNRRSDSAGTSSGPGCEDDVSCNOFNBOI 562
 QY 536 CSG-HGQSCGDCICLDSWMTGYNCNCTTRDTQSSNGLLCSGRKCEGSCVCIQPGS 593
 DB 563 CGSPERKGNCKGCKCKLIGNCTACELSMEKCTTDDGLVGNHGNCTGCKVC--ENQ 620
 QY 594 Y-GDICEKCPCTPDACITRKCEVCKKFRDGLADHENTCNRYCDEI--ESVRELDTG 649
 DB 621 YTGFECEQPCAPCDKCLEYTPCVOCKAFKFLQER-CLRECKMEIFHDKIRE----G 675
 QY 650 KDAVCTYKNEDDVCVRRQYEDSSGKIIIVVEPECKRPDILVULSMAGAILLGL 709
 DB 676 HGICLPKIDDECAVITTYEYDRGKVIVIAQITKVCDDGLNLAIVGAGVAVGL 735
 QY 710 AALLIWLKLLITIHDKERAFKEERARAKMDTANPLKREATSTFTNTITY 759
 DB 726 FULLIMKLLTFIHDTREFAKFEKRNANKMDTNGENPIYKATSTFKNPTY 785

RESULT 14

ID Q9NAST7 PRELIMINARY; PRT; 837 AA.

AC Q9NAST7;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)

DE INTEGRIN BETA SUBUNIT PRECURSOR.

GN BINT.

OS Anopheles gambiae (African malaria mosquito).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;

OC Anophelinae.

NCBI_TaxID=7165;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-G3;

RA MEDLINE=2131701; PubMed=11437913;

RT "Beta-Integrin of Anopheles gambiae: mRNA cloning and analysis of

RT structure and expression."

RL Insect Mol. Biol. 10:217-223(2001).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE

CC BONDS (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.

DR EMBL: AJ292755; CAC00630.1; -

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR002369; Integrin_B.

DR InterPro: IPR001169; Integrin_beta_C.

DR InterPro: IPR003659; PSI.

DR InterPro: IPR002035; VWEA.

DR Pfam: PF00362; Integrin_B_1.

DR PRINTS: PR01186; INTEGRINB.

DR ProDom: PD001811; Integrin_B_1.

DR SMART: SM00187; INB.1.

DR SMART: SM00423; PSI.1.

DR SMART: SM00327; VWA.1.

DR PROSITE: PS00022; EGF_1; UNKNOWN_2.

DR PROSITE: PS01186; EGF_2; UNKNOWN_1.

DR PROSITE: PS00243; INTEGRIN_BETA_3.

KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;

KW Integrin; Repeat; Signal; Transmembrane.

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

FT SIGNAL

Query Match 38.7%; Score 1609; DB 5; Length 837;

Best Local Similarity 39.5%; Pred. No. 5.6e-125; Mismatches 284; Indels 72; Gaps 19;

Tue May 21 14:53:25 2002

us-09-673-302a-1.rsp

Page 12

[illegible]

Search completed: May 19, 2002, 12:26:43
Job time: 6465 sec